THE JOURNAL OF GENERAL PSYCHOLOGY

Experimental, Theoretical Clinical and Historical Psychology

EDITED BY
CARL MURCHISON

FRANK ANGELL
Stanford University
F. C. BARTLETT
Cambridge University
V. M. BOROVSKI
Gosudarstvennei Institut Eksperimentalnoi Psikhologii, Moskva
G. S. BRETT
University of Toronto
KARL BÜHLER
Universität, Wien
HARVEY CARR
University of Chicago
W. J. CROZIER
Harvard University
RAYMOND DODGE
Yale University
JAMES DREVER
University of Edinburgh
GEORGES DUMAS
Université de Paris
KNIGHT DUNLAP
The Johns Hopkins University
SHEPHERD IVORY FRANZ
University of California at Los Angeles
H. L. HOLLINGWORTH
Columbia University
T. A. HUNTER
Victoria University College, New Zealand
PIERRE JANET
Université de Paris
F. KIESOW
Regia Università degli Studi Torino
OTTO KLEMM
Universität, Leipzig
FELIX KRUEGER
Universität, Leipzig
WILLIAM MCDOUGALL
Duke University
A. MICHOTTE
Université de Louvain
WALTER MILES
Stanford University
JOHN PAUL NAFE
Clark University
A. L. SCHNIERMANN
Bekhterev Refleksologicheskii Institut po Izucheniya Mozga, Leningrad
C. E. SPEARMAN
University of London
EDWARD C. TOLMAN
University of California
JOHN B. WATSON
New York City
ALBERT P. WEISS
Ohio State University
H. P. WELD
Cornell University
RAYMOND H. WHEELER
University of Kansas
R. W. WILCOCKS
Uniwersiteit van Stellenbosch, Suidafrika
CLARK WISSLER
Yale University
ROBERT M. YERKES
Yale University

LUBERTA M. HARDEN, PH.D.
Assistant Editor

VOLUME IV

1 9 3 0

Published quarterly by
CLARK UNIVERSITY PRESS
Worcester, Massachusetts
U. S. A.

Copyright, 1930
by
CLARK UNIVERSITY

34-127
12-7
6

TABLE OF CONTENTS

ARTICLES

LOY W. ALLISON AND JOSEPH PETERSON
Effects of Visual Exposure on the Rate and Reliability of Stylus-Maze Learning 36-48

T. CUNLIFFE BARNES AND B. F. SKINNER
The Progressive Increase in the Geotropic Response of the Ant *Aphaenogaster* 102-112

F. C. BARTLETT
Experimental Method in Psychology 49-66

J. E. CASTER
Emotional Reactions to Strong Stimuli 131-153

RAYMOND DODGE
Fundamental Steps in the Development of Adaptive Behavior of the Eyes 3-14

C. E. FERREE AND GERTRUDE RAND
Size of Stimulus and Brightness of Preexposure and of Surrounding Field in Relation to the Extent and Shape of the Color Fields 15-35

G. L. FREEMAN
Changes in Tonus During Completed and Interrupted Mental Work 309-334

KATE HEVNER
An Empirical Study of Three Psychophysical Methods . . 191-212

DOROTHEA E. JOHANNSEN
A Quantitative Study of Binocular Color Vision . . . 282-308

HELEN L. KOCH
The Influence of Some Affective Factors Upon Recall . . 171-190

GEORGE KREEZER
Luminous Appearances 247-281

JOHN A. MCGEOCH AND VERA J. OBERSCHELP
The Influence of Length of Problem and of Transfer upon Rational Learning and its Retention 154-170

VERA J. OBERSCHELP AND JOHN A. MCGEOCH
The Influence of Length of Problem and of Transfer upon Rational Learning and its Retention 154-170

JOSEPH PETERSON AND LOY W. ALLISON
Effects of Visual Exposure on the Rate and Reliability of Stylus-Maze Learning 36-48

GERTRUDE RAND AND C. E. FERREE
Size of Stimulus and Brightness of Preexposure and of Surrounding Field in Relation to the Extent and Shape of the Color Fields 15-35

HAROLD D. SCOTT
Hypnosis and the Conditioned Reflex 113-130

B. F. SKINNER AND T. CUNLIFFE BARNES
The Progressive Increase in the Geotropic Response of the Ant *Aphaenogaster* 102-112

T. J. B. STIER
"Spontaneous Activity" of Mice 67-101

CORNELIA D. TAYLOR
Visual Perception versus Visual Plus Kinaesthetic Perception in Judging Colored Weights 229-246
R. H. WATERS
The Influence of Large Amounts of Manual Guidance upon Human Maze Learning 213-228

SHORT ARTICLES AND NOTES

J. CARVER ANDERSON AND PAUL L. WHITELY
The Influence of Two Different Interpolations upon Time Estimation 391-401
A. F. CHAISSON
An Alternative Approach to the Mathematical Study of Learning Curves 352-359
C. E. COGHILL
The Development of Half-Centers in Relation to the Question of Antagonism in Reflexes 335-338
E. W. CUMMINS AND JOHN M. STALNAKER
A Quantitative Method of Testing the Diminution of the Effect of Post-Hypnotic Suggestion with the Passage of Time 366-372
GEORGE F. DUNKELBERGER AND ELDON K. RUMBERGER
An Experimental Study in Perceptual Insight . . . 385-389
MARTIN F. FRITZ
Experimental Evidence in Support of Professor Thurstone's Criticism of the Phi Gamma Hypothesis 346-352
CARL MURCHISON
Notice of Special Volume 335
DAVID W. OATES
The Energy Concept in Modern Psychology 401-406
MARION W. RICHARDSON AND JOHN M. STALNAKER
Time Estimation in the Hypnotic Trance 362-366
A. L. SCHNIERMANN
Second Anniversary of the Death of Bekhterev . . . 391
B. F. SKINNER
On the Inheritance of Maze Behavior 342-346
JOHN M. STALNAKER AND E. W. CUMMINS
A Quantitative Method of Testing the Diminution of the Effect of Post-Hypnotic Suggestion with the Passage of Time 366-372
JOHN M. STALNAKER AND MARION W. RICHARDSON
Time Estimation in the Hypnotic Trance 362-366
EDWARD G. STOY
A Preliminary Study of Ocular Attitudes in Thinking of Spatial Relations 379-385
EDWARD CHACE TOLMAN
Maze Performance a Function of Motivation and of Reward as Well as of Knowledge of the Maze Paths . . . 338-342
WILLARD L. VALENTINE
A Study of Learning Curves: II. The Relationship between the Hyperbola and the Arc Cotangent Function . . . 359-362
PAUL L. WHITELY AND J. CARVER ANDERSON
The Influence of Two Different Interpolations upon Time Estimation 391-401
HELEN MORRILL WOLFLE
Time Factors in Conditioning Finger Withdrawal . . 372-378

APPARATUS

CHESTER W. DARROW
Continuous Quantitative Records in Standard Units by the Wheatstone Bridge Method 418-420
PAUL R. FARNSWORTH
Bell Laboratory Victor Records 415-416
G. C. GRINDLEY
Apparatus for a Group Demonstration of Weber's Law . 427-429
STARKE ROSECRANS HATHAWAY
A Pendulum Chronoscope 423-427
MAX F. MEYER AND WILLARD L. VALENTINE
A Description of the Lectometer and the Reliability of Lectometer Scores 407-415
FRANKLIN C. PASCHAL
An Interval Timer 420-423
ROBERT H. SEASHORE
Phonograph Records from Bell Telephone Laboratories Illustrating Phenomena of Acoustics 416-417
WILLARD L. VALENTINE AND MAX F. MEYER
A Description of the Lectometer and the Reliability of Lectometer Scores 407-415

BOOKS

W. KÖHLER
Gestalt Psychology (R. M. Ogden) 430-434
W. M. MARSTON
Emotions of Normal People (Frank A. Geldard) . . . 438-441
M. PRINCE
Clinical and Experimental Studies in Personality (Edmund S. Conklin) 434-435
R. H. WHEELER
The Science of Psychology (P. H. Ewert) 435-438

$7.00 per volume
Single numbers $2.00
This issue $7.00

QUARTERLY

December, 1930
Volume IV, Numbers 1-4

Founded in 1927 by Edward Bradford Titchener and Carl Murchison

THE JOURNAL OF GENERAL PSYCHOLOGY

Experimental, Theoretical, Clinical, and Historical Psychology

---

SPECIAL VOLUME ISSUED IN DECEMBER, 1930

Chaque article est suivi d'un résumé en français
Jedem Artikel wird ein Referat auf deutsch folgen

FUNDAMENTAL STEPS IN THE DEVELOPMENT OF ADAPTIVE BEHAVIOR OF THE EYES 3
By Raymond Dodge
SIZE OF STIMULUS AND BRIGHTNESS OF PREEXPOSURE AND OF SURROUNDING FIELD IN RELATION TO THE EXTENT AND SHAPE OF THE COLOR FIELDS . . . 15
By C. E. Ferree and Gertrude Rand
EFFECTS OF VISUAL EXPOSURE ON THE RATE AND RELIABILITY OF STYLUS-MAZE LEARNING 36
By Joseph Peterson and Loy W. Allison
EXPERIMENTAL METHOD IN PSYCHOLOGY 49
By F. C. Bartlett
"SPONTANEOUS ACTIVITY" OF MICE 67
By T. J. B. Stier
THE PROGRESSIVE INCREASE IN THE GEOTROPIC RESPONSE OF THE ANT *APHAENOGASTER* 102
By T. Cunliffe Barnes and B. F. Skinner
HYPNOSIS AND THE CONDITIONED REFLEX 113
By Harold D. Scott
EMOTIONAL REACTIONS TO STRONG STIMULI 131
By J. E. Caster
THE INFLUENCE OF LENGTH OF PROBLEM AND OF TRANSFER UPON RATIONAL LEARNING AND ITS RETENTION 154
By John A. McGeoch and Vera J. Oberschelp

Worcester, Massachusetts
Copyright, 1930, by Clark University
Entered as second-class matter December 1, 1927, at the post-office at Worcester, Mass., under the Act of March 3, 1879

CONTENTS (*continued*)

THE INFLUENCE OF SOME AFFECTIVE FACTORS UPON RECALL 171
By Helen L. Koch

AN EMPIRICAL STUDY OF THREE PSYCHOPHYSICAL METHODS 191
By Kate Hevner

THE INFLUENCE OF LARGE AMOUNTS OF MANUAL GUIDANCE UPON HUMAN MAZE LEARNING . . . 213
By R. H. Waters

VISUAL PERCEPTION VERSUS VISUAL PLUS KINAESTHETIC PERCEPTION IN JUDGING COLORED WEIGHTS 229
By Cornelia D. Taylor

LUMINOUS APPEARANCES 247
By George Kreezer

A QUANTITATIVE STUDY OF BINOCULAR COLOR VISION 282
By Dorothea E. Johannsen

CHANGES IN TONUS DURING COMPLETED AND INTERRUPTED MENTAL WORK 309
By G. L. Freeman

SHORT ARTICLES AND NOTES 335

- Notice of Special Volume (Carl Murchison) 335
- The Development of Half-Centers in Relation to the Question of Antagonism in Reflexes (C. E. Coghill) 335
- Maze Performance a Function of Motivation and of Reward as Well as of Knowledge of the Maze Paths (Edward Chace Tolman) 338
- On the Inheritance of Maze Behavior (B. F. Skinner) . . . 342
- Experimental Evidence in Support of Professor Thurstone's Criticism of the Phi Gamma Hypothesis (Martin F. Fritz) 346
- An Alternative Approach to the Mathematical Study of Learning Curves (A. F. Chaisson) 352
- A Study of Learning Curves: II. The Relationship between the Hyperbola and the Arc Cotangent Function (Willard L. Valentine) 359
- Time Estimation in the Hypnotic Trance (John M. Stalnaker and Marion W. Richardson) 362
- A Quantitative Method of Testing the Diminution of the Effect of Post-Hypnotic Suggestion with the Passage of Time (John M. Stalnaker and E. W. Cummins) 366
- Time Factors in Conditioning Finger Withdrawal (Helen Morrill Wolfle) 372
- A Preliminary Study of Ocular Attitudes in Thinking of Spatial Relations (Edward G. Stoy) 379
- An Experimental Study in Perceptual Insight (George F. Dunkelberger and Eldon K. Rumberger) 385
- Second Anniversary of the Death of Bekhterev (A. L. Schniermann) 390
- The Influence of Two Different Interpolations upon Time Estimation (Paul L. Whitely and J. Carver Anderson) . . . 391
- The Energy Concept in Modern Psychology (David W. Oates) 401

CONTENTS (*continued*)

APPARATUS 407
A Description of the Lectometer and the Reliability of Lectometer Scores (Willard L. Valentine and Max F. Meyer) . 407
Bell Laboratory Victor Records (Paul R. Farnsworth) . . 415
Phonograph Records from Bell Telephone Laboratories Illustrating Phenomena of Acoustics (Robert H. Seashore) . . 416
Continuous Quantitative Records in Standard Units by the Wheatstone Bridge Method (Chester W. Darrow) . . 418
An Interval Timer (Franklin C. Paschal) 420
A Pendulum Chronoscope (Starke Rosecrans Hathaway) . . 423
Apparatus for a Group Demonstration of Weber's Law (G. C. Grindley) 427
BOOKS 430
W. Köhler, "Gestalt Psychology" (R. M. Ogden) 430
Other Books 434

FUNDAMENTAL STEPS IN THE DEVELOPMENT OF ADAPTIVE BEHAVIOR OF THE EYES*[1]

From the Institute of Human Relations, Yale University

Raymond Dodge

Summary

1) Records of human eye-movements involved in following a simple object of regard during its harmonic oscillation, in the foveal fixation of a motionless object, and in the fixation movements to new objects of regard, show characteristics which lead us to the following generalizations and hypotheses:

2) Approximation and correction is a type of adjustment that appears in simple forms of human behavior as well as in the conduct of affairs and the development of science. While the first responses to a new situation often follow the principle of trial and selection, subsequent adjustment in the selected mode of response commonly follows other formulae, frequently that of approximation and correction. The latter seems to be a complement of the former principle. We conjecture that it ranks close to trial and error as a general formula of adjustment.

3) The tendency of an adjustment pattern to break down at or near its physiological limit and to give place to a simpler form of reaction appears in response to relatively simple stimuli as well as in response to the complex stimuli of our social and economic environment. We conjecture that it represents some fundamental characteristic of neural action that may be found in a great variety of human and infra-human behavior. It is probably related to relative fatigue and to the rivalry and competition for control of common paths.

4) Anticipatory reaction is found in a simple form of human adaptive behavior as well as that which depends on insight and purpose. We conjecture that it is one of the fundamental forms of learning to meet familiar situations. Whether it develops on the

---

*Received in the Editorial Office, June 10, 1929.

[1] With grateful acknowledgments to my collaborators, Dr. R. C. Travis, Research Associate in the Institute of Psychology, and Dr. J. C. Fox, Clinical Assistant Professor of Neurology in the Yale School of Medicine.

basis of antecedent cues which are so related to the adequate stimuli that they become secondary or conditioned stimuli, or on the basis of the constellation of stimuli in configurations, is a question that our data did not answer.

Introduction

For purposes of this report which relates primarily to pursuit eye-movements, let us assume an instrument which permits the photographic registration of conjugate eye-movements in response to the harmonic oscillation of a simple object of regard. The main facts of the development and breakdown of the relevant ocular behavior patterns may then be presented without technical prelude by reproducing typical photographic records.

Figures 1 to 4 show the eye-movements in the attempt to follow a simple object in harmonic oscillation. Since they were taken on strips of sensitized paper which moved downwards past a horizontal exposure slit, they should all be read from the bottom upwards. Time is indicated in tenths of a second on the right of each record. The double white lines show the oscillation of the fixation object moving through 6 degrees of the visual field, to the right and left of the point of regard. Line *H* shows the movements of the head which was stabilized by mouth-bit and two rests near the back. With these rests there was very little head movement and the record is consequently approximately straight. Line *E* shows the conjugate movements of a mechanically closed eye, while the other (open) eye followed the oscillating object. At frequencies of one oscillation of the object in 1.5 or more seconds, the eyes of all subjects followed the moving object quite well after the first half oscillation, as shown in Figures 1 and 2. At higher frequencies pursuit was less regular and showed marked individual differences. The amplitudes of the several curves are not directly comparable apart from routine controls which are not shown in the figures.

The Development of an Adequate Adjustment Pattern

In most of our records the first half-oscillation presents a very characteristic picture, which is illustrated by Figures 1 and 2. After the object begins to move, the recording eye stands approximately still for a time, corresponding to a latency of the order of 0.2″. Then it makes an ineffectual corrective jump in the direction of objective movement, which passes into a pursuit glide. The corrective jump is an eye-movement of the saccadic type such as usually corrects a

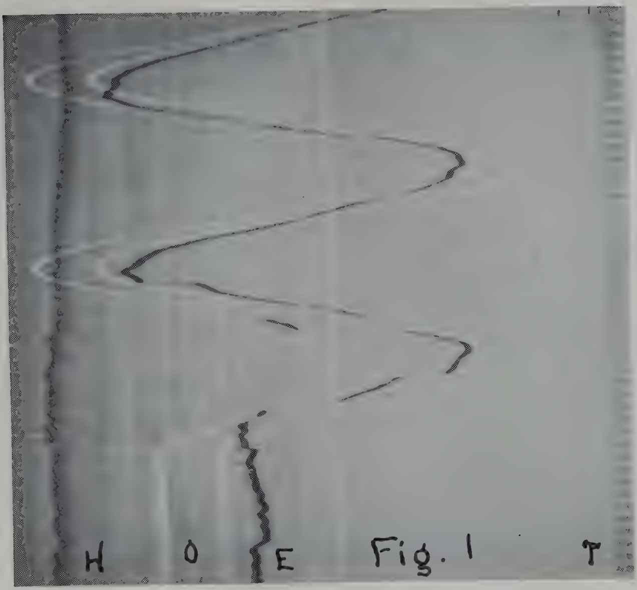


Fig. 1

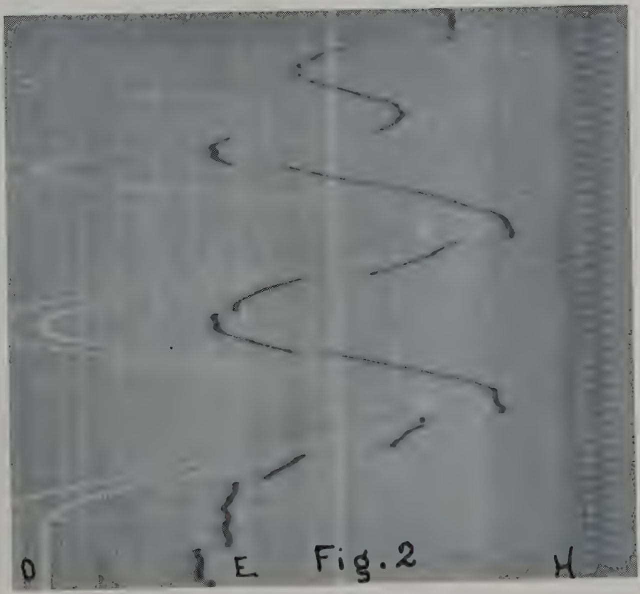


Fig. 2

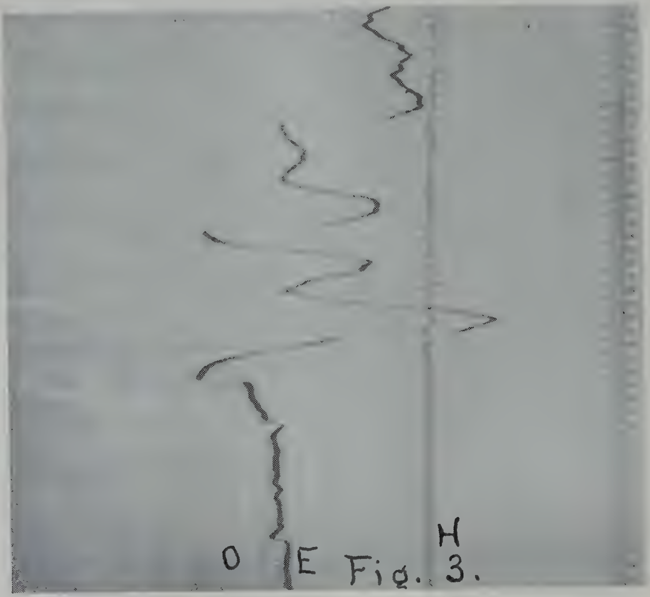


Fig. 3.

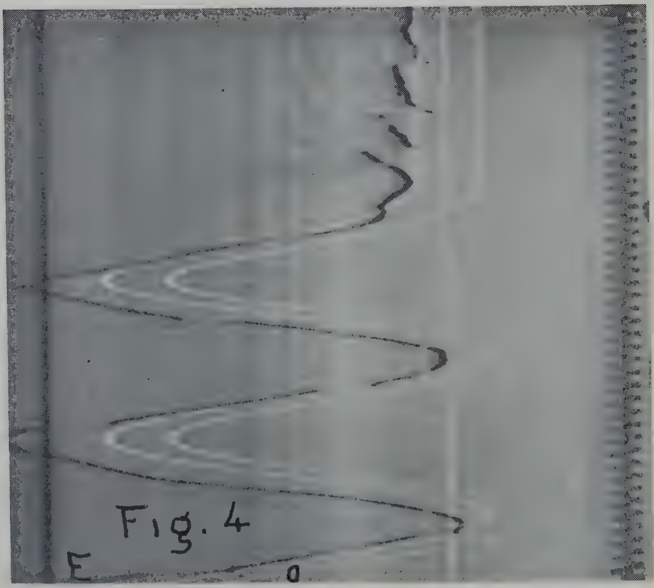


Fig. 4

lapsed fixation or the new fixation of a peripheral object. The glide is the first indication of true pursuit. The slope of the first glide does not correspond to the contemporary velocity of the object at the moment of response, but to an earlier and lesser velocity. Both the saccadic movement and the first pursuit glide show the elapse of a latent time between the movement of the object and the response of the eye. This contrasts sharply with later phases of pursuit where latent time seems to disappear in favor of anticipatory reactions.

Figure 3 shows an important variant of the general picture of the first phase of pursuit. Possibly due to some more or less clear anticipation of which the record is the only indicator, the eye-line shows a preliminary slow glide to the right before the first saccadic refixation. In this record the preliminary glide is almost simultaneous with the beginning of objective motion in the other direction. It is obviously an anticipatory reaction in the wrong direction. This is followed by an approximately normal beginning reaction consisting of a saccadic movement in the right direction (of normal latency) which passes into a particularly inadequate pursuit and a second, long saccadic movement to the extreme point of oscillation. The unique feature of the record is the anticipatory false start. If the start had been made in the right direction, this initial glide might have been read by the uninitiated as a true reaction with an extraordinarily short latency. Such records should make us very conservative in measuring latencies from initial glides. This exceptional record only serves to emphasize the general pattern of the first phase with its saccadic correction of lapsed fixation and its ineffectual glides, both after obvious reaction-times.

This pattern changes notably in the second phase of oscillation. Whenever the objective movement is not too rapid all phases after the first show a much closer approximation of the ocular pursuit to the actual motion of the object, while latency is apparently decreased or eliminated. Corrective movements, however, are never entirely lacking. The appearance of adequate ocular pursuit after the first phase of inadequate pursuit seems to represent a learning process of extraordinarily rapid development.

The percentage amplitudes of pure pursuit glides at different oscillation frequencies are shown in Figure 5. The data for this figure were obtained from the eye-movement records of eight subjects pursuing a small visual object in harmonic oscillation. Each point in the curve represents the average percentage amplitudes of

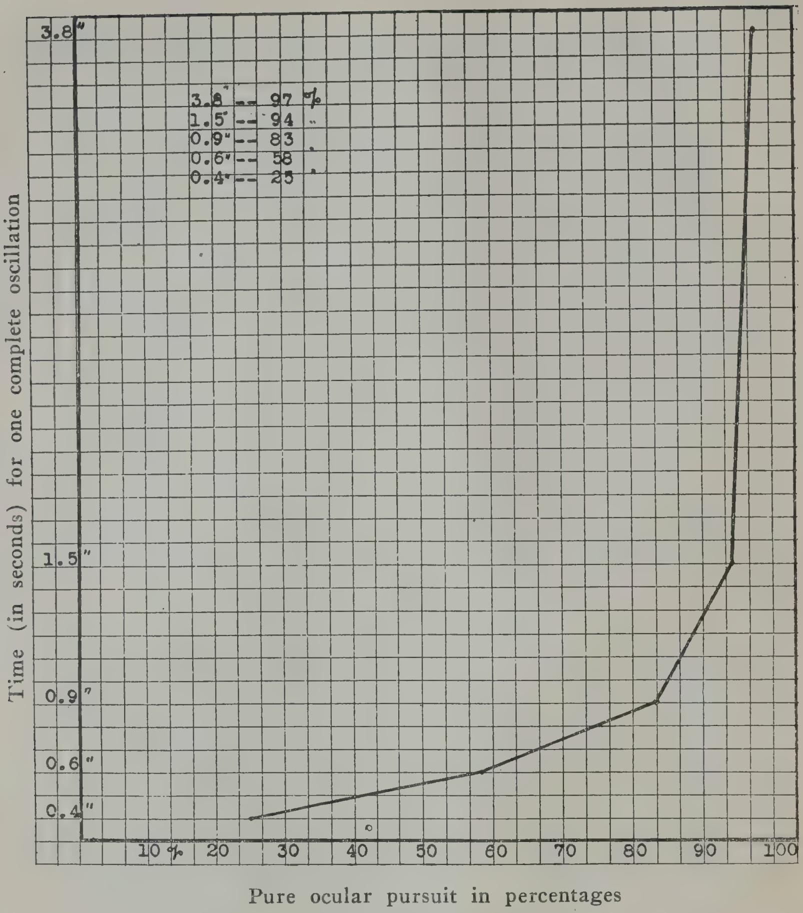


FIGURE 5

Relation between Pursuit Adequacy and Speed of Object
Each point represents the average of 400 oscillations for eight observers

pure pursuit glides during 400 phases of oscillation. The form of the curve indicates that 100% of pure pursuit will never be reached at any oscillation frequency, while zero pursuit may be expected at slightly lower frequencies than those actually used.

The theoretical implications of these records are clear enough. Even such a simple and practiced act as the ocular pursuit of an object in harmonic oscillation does not start perfectly formed in successive instances, but shows a temporal development of approximation and correction in each case. As the adjustment becomes more and more nearly adequate there is a consequent increase of true pursuit up to a limit of approximation which differs with each frequency of oscillation. Corrective saccadic eye-movements, however, follow a quite different course. In the first phase of movement they vary irregularly with the oscillation frequency. They are most extensive at the beginning of adjustment, and in later phases at moderately high frequencies. They are especially long near the threshold of the break after the true pursuit becomes notably inadequate, but disappear entirely before the pursuit glides disappear.

Approximative and corrective eye-movements are not peculiar to ocular pursuit. They have been found in various photographic records of eye-movements. As was pointed out first by Dearborn and then by various others, objects which are first seen peripherally are commonly fixated in from two or three saccadic jerks. The first movement in the direction of the peripheral object is a more or less gross approximation which is corrected by a second and sometimes even by a third readjustment. Moreover, even the fixation of a motionless object is never maintained absolutely accurately. Photographic records always show a more or less close approximation to fixation with corrective refixation movements at irregular intervals.

Much overt human behavior and human thinking seem to follow a pattern in which a rough approximation is followed by a phase of correction, which in turn may be only another approximation. I do not know how general this pattern is, but I conjecture that it is a common biological phenomenon and an almost universal way of meeting situations after the first crude trial and selection gives the general pattern of response. Our practical adjustments commonly follow this formula, from driving an automobile to hanging a picture—from lining up a lens system to scientific writing. As I understand it, even science progresses according to this formula, whether it deals with astral physics or atomic weights. In such cases, however, the actual procedure may be exceedingly difficult to record and analyze. The main point is not an effort to extend the formula, but to show that a formula which appears in so much of our thinking and behavior appears also in the development of such simple acts as the eye-movements of pursuit and fixation where it may be adequately recorded and studied.

THE BREAKDOWN OF ADEQUATE OCULAR PURSUIT ADJUSTMENT

Figure 2 shows the movement of an eye during accelerating oscillation of the object of regard. When the objective movement becomes too rapid, the amplitude of pure pursuit decreases, while the saccadic corrective movements become longer and more numerous. As the oscillation accelerates still further, both the corrective movements and the pursuit tend to disappear into approximately still fixation. This is more clearly seen in Figure 3 where one oscillation takes about 0.4″. During the first part of this record the pure pursuit movements are relatively short and the corrective movements long. At the end of the record pursuit practically breaks down, the long corrective movements disappear and approximately still fixation supervenes. In addition to indicating the limits to adequate pursuit, these records, Figures 2 and 3, give some important information concerning the breakdown of an adequate adjustment pattern as the objective situation becomes too difficult.

Normal and abnormal psychology seems to furnish innumerable instances of analogous breaks and the substitution of simpler reaction patterns. When situations entirely preclude adequate adjustment, whether in our day's work, in social contacts, or in reactions to economic pressure, inadequate behavior patterns do not persist indefinitely. In response to some unknown neural mechanism, possibly due to relative fatigue, possibly to the law of effect or to some other condition of retroflex action, the inadequate pattern is abandoned, while another and in many cases a simpler reaction pattern more or less analogous to still fixation of the eyes takes its place. Under such circumstances the human reactor may rest, take a vacation, leave the room, arrange for marital separation, go into bankruptcy, or take refuge in the world of fantasy. Certainly a great many human breakdowns seem to follow the general formula that is found in the simple pursuit reactions of the eyes when the objective movement of the object of regard becomes too complicated.

Our second point is that in the breakdown of adjustive pursuit eye-movements we find an analogy to a wide variety of breakdowns of human adjustment.

Anticipatory Reaction as a Form of Learning

Still another feature of our eye-movement records deserves consideration. As we have already noted, the development of approximately adequate ocular adjustment to moving objects involves the reduction or elimination of reaction latency. After the first phase of pursuit the slowing down and speeding up of the eye-movements does not wait for the slowing down and speeding up of the object with an intervening latent time. In many records the eyes actually start their return sweeps coincidentally with the object, or even before it. In terms of traditional psychology this constitutes what we commonly call an anticipatory reaction. The facts may be seen clearly in the more adequate pursuit adjustments of Figure 1, where, after the first phase of oscillation, the changes in the direction of eye-movement are either coincident with, or antecedent to, changes in the direction of the movement of the object. The disappearance of latency is doubtless only apparent. It would be the least plausible of the various hypotheses which we have considered to assume a short circuiting or automatizing process by which the latency was suddenly reduced to the point of disappearance. The crucial evidence against such an hypothesis would be two aspects of our records, one of which we have already noted, namely, that the reaction may antedate the original stimulus and the eye may turn before the object reaches the extreme point of this swing. This seems to preclude a short-circuiting process. The second evidence against this hypothesis is that when the object comes to a sudden stop, eye-movements may continue for an appreciable time, as is shown in Figure 4.

Two somewhat more plausible hypotheses have been considered to account for the apparent loss of latency. It might be explained as a case of anticipatory reaction in which eye-movement followed cues which were associated with the original stimulus. In that case one would regard the pursuit eye-movements as a kind of conditioned response in which a given velocity of the object evokes a very different velocity of eye-movement.

The first argument against this hypothesis is the extreme complexity of the reaction system which it would involve, together with the rapidity with which each case of pursuit becomes adequate. The second argument against regarding adequate pursuit movements as a chain of conditioned reactions is that any specific angular velocity of the object occurs twice in each phase of oscillation, once in acceleration and once in deceleration. A conditioned eye-movement evoked by any particular speed of the object would consequently depend on some clue furnished by the general configuration of the oscillation before it could respond in smooth pursuit to the two different meanings of any angular velocity of the object. That is to say, no speed clue would be adequate to determine a conditioned pursuit reaction independent of its setting in the oscillation pattern.

This leads us to a second hypothesis, namely, that the objective motion is apprehended as a unitary acceleration pattern to which the eye reacts in a unitary pattern of response. This hypothesis lays especial emphasis on the approximately smooth pursuit movements of the eye after the first phase. Such smooth pursuit seems irreconcilable with a chain of conditioned reactions. Its peculiar sine-wave character suggests some kind of integration into a unitary response. The conditions of that integration would seem to be given in the character of the stimulus situation as apprehended during the first half-oscillation. At any rate, the difference between the first and successive phases of pursuit forces us to regard the latter as a product of a learning process.

Further evidence in support of this hypothesis is furnished by the regularity of some of the inadequate pursuit patterns and especially by the character of the post-excitation pattern after the sudden cessation of objective motion as seen at the top of Figure 4.

That this hypothesis is only partially satisfactory is indicated by the small corrective movements that occur even in the smoothest pursuit. These are of various amplitudes, sometimes positive and sometimes negative in direction. They are in all probability specific reactions to lapsed fixation. Moreover, pursuit may falter at speeds well below the actual capacity of the eyes to move and may include on occasion long and irregular saccadic corrections, with occasional moments of adequate pursuit. These facts would seem to be inconsistent with the closed integration of behavior patterns.

A fourth possibility should not be overlooked. It is conceivable that both the second and the third hypotheses might be combined and that the corrections represent specific reactions to lapsed fixations which are superposed on a more or less adequately integrated pattern Certainly the second and third hypotheses are not mutually exclusive. Both might interact and support each other. At present the easiest reconstruction to account for all of the facts seems to be the hypothesis of occasional cues superposed on total situation patterns. This falls in line with various casual and experimental observations of the interrelation between general form and detail.

Whether we regard an ocular pursuit as a chain of conditioned responses or an integrated response to configurated situations, it remains clear that we have to do with some form of anticipatory reaction, which is represented in many of our practical responses. Intelligent persons do not wait for events to occur before they adjust to them, whether they are collisions with approaching automobiles, economic issues, social functions, political responsibilities, old age, or death. The systematic sequences in all these events furnish preliminary cues for anticipatory reactions. Here again we would not generalize from the eye-movement. The main point is that the reactions of ocular pursuit involve anticipatory adaptative responses which may be accurately recorded and analyzed and that these responses seem to be analogous to a great variety of other learned adjustments which it is difficult or impossible to study with the same precision.

As has appeared in this discussion, the important generalization is not from simple to complex behavior, not even that simple formulae which are found in the eye-movements apply to a multitude of miscellaneous adjustments, but rather that formulae of apparently widespread application to human behavior can be studied and recorded in such relatively simple phenomena as the movements of the eyes.

Yale University
New Haven, Connecticut

LES ÉTAPES FONDAMENTALES DU DÉVELOPPEMENT DE L'ADAPTATION DES YEUX

(Résumé)

Les résultats des études des mouvements des yeux chez les êtres humains quand ils suivent du regard un simple objet pendant son oscillation harmonique, dans la fixation fovéale d'un objet immobile, et dans les mouvements de fixation aux nouveaux objets du regard, montrent des caractéristiques qui nous font formuler les généralisations et les hypothèses suivantes: L'approximation et la correction sont un type d'ajustement lequel se trouve dans les simples formes du comportement humain ainsi que dans la conduite des affaires et le développement de la science. Tandis que les premières réponses à une situation nouvelle suivent souvent le principe essai et choix, l'ajustement subséquent dans la mode choisie de réponse suit ordinairement d'autres formules, fréquemment celle d'approximation et de correction. Celle-ci semble être un complément du premier principe. On conjecture que celui-ci a à peu près la valeur du principe essai et erreur comme formule générale d'ajustement. La tendance d'une mode d'ajustement à disparaître à ou près de sa limite physiologique et à donner lieu à une plus simple forme de réaction se montre en réponse aux stimuli relativement simples ainsi qu'en réponse aux stimuli complexes de notre milieu social et économique. On conjecture qu'elle représente quelque trait fondamental de l'action des nerfs lequel peut se trouver dans une grande variété du comportement humain et infra-humain. Elle aura des rapports avec la fatigue relative et la rivalité et la concurrence pour le contrôle des sentiers communs. La réaction anticipée se trouve dans une simple forme du comportement d'adaptation des humains ainsi que dans celle qui dépend

de connaissance et de but. On conjecture que c'est une des formes fondamentales de l'effort d'apprendre à faire face à des situations bien connues. Nos données ne nous disent pas définitivement si elle se développe sur la base de suggestions antérieures lesquelles ont tel rapport avec les stimuli adéquats qu'elles deviennent des stimuli secondaires ou conditionnels ou sur la base de la position des stimuli dans les configurations.

DODGE

WESENTLICHE STUFEN IN DER ENTWICKLUNG DES ANPASSUNGSVERMÖGENS DER AUGEN

(Referat)

Aufzeichnungen von Bewegungen menschlicher Augen, während ihrer harmonischen Schwingungen beim Verfolgen eines einfachen Sehobjektes, bei fovealer Fixierung eines unbeweglichen Objektes, und bei Fixierungbewegungen auf neue Sehobjekte, weisen Merkmale auf, die uns zu den folgenden Verallgemeinerungen und Hypothesen führen: Annäherung (approximation) und Korrektur ist ein Anpassungstypus, welcher sowohl in einfachen Formen des menschlichen Verhaltens als auch in der Durchführung von Geschäften und in der Entwicklung der Wissenschaften wiederkehrt. Währenddem die ersten Reaktionen hinsichtlich einer Situation oft dem Prinzip von Versuch und Auswahl, folgen nachträgliche Anpassungen im gewählten Modus der Reaktion gewöhnlich andern Formeln, oft der der Annäherung und Korrektur. Das letztere Prinzip scheint eine Ergänzung des obigen zu sein. Wir vermuten, dass es dem Probieren als allgemeine Formel der Anpassung nahesteht. Die Tendenz eines Anpassungsmusters, an oder nahe einer Physiologischen Grenze zu zerfliessen, um einer einfacheren Form der Reaktion Platz zu machen, zeigt sich sowohl bei der Rückwirkung auf verhältnismässig einfache Reize, wie auch bei der Rückwirkung auf die komplexen Reize unserer gesellschaftlichen und ökonomischen Umgebung. Wir vermuten, dass dies ein wesentliches Merkmal der Nerventätigkeit ist, das wohl in manigfachen Variationen des menschlichen und inframenschlichen Verhaltens wiederzufinden ist. Es ist wahrscheinlich verwandt mit der relativen Müdigkeit und dem Wetteifer und—Kampf für die Beherrschung allgemeiner Mittel. Zuvorkommende Reaktionen kann man bei einfachen Formen des menschlichen Anpassungsvermögens beobachten, als auch bei solchen, die auf Einsicht und Ziel beruhen. Wir vermuten, dass sie eine der wesentlichen Formen des Lernens sind, um bekannten Situationen zu begegnen. Ob sie sich auf Grund vorhergehender Winke entwickeln, die sich zu den ihnen entsprechenden so verhalten, dass sie sekundäre oder bedingte Reize werden, oder auf Grund einer Reizkonstellation in Gebilden, ist eine Frage, die unsere Aufzeichnungen nicht bestimmt beantworten.

DODGE

SIZE OF STIMULUS AND BRIGHTNESS OF PREEXPOSURE AND OF SURROUNDING FIELD IN RELATION TO THE EXTENT AND SHAPE OF THE COLOR FIELDS*

From the Research Laboratory of Physiological Optics, Wilmer Ophthalmological Institute, Johns Hopkins Medical School

C. E. Ferree and Gertrude Rand

In determining the influence of a variable factor the customary and approved procedure is to hold constant at some predetermined value all factors but the one whose effects are under investigation. A very incomplete picture of the effect of the factor is had, however, if the investigation stops here. That is, very different results are obtained when different values are selected for one or more of the factors which are to be held constant during the series.

In previous papers (1-4) we have shown the effect of intensity of stimulus and brightness of preexposure and surrounding field on the extent and shape of the color fields. No systematic attempt was made, however, to show the interaction of these effects. It is the purpose of the present paper to show the effect of size of stimulus and to include also the interaction of effect of size of stimulus and brightness of preexposure and surrounding field. That is, in this paper the effect of size of stimulus will be shown for different brightnesses of preexposure and surrounding field, and conversely the effect of brightness of preexposure and surrounding field will be shown for different sizes of stimulus. For a more complete understanding of the range and extent of the influence of the variable factors and their interacting effects, similar studies should be made for intensity and size of stimulus, intensity of stimulus and brightness of preexposure and surrounding field, size of stimulus and state of adaptation of the retina, intensity of stimulus and brightness of preexposure and surrounding field, size of stimulus and state of adaptation of the retina, etc. In this connection it will be remembered that complications of effect even greater than those suggested here character-

---

*Received for publication by Carl Murchison of the Editorial Board, May 16, 1929.

ize the conditions of work in perimetry unless great care is exercised in the control of variable factors.

Three possibilities may be noted as to the effect of increase in the size of the stimulus on the response of a sense organ. (*a*) There is an increase in the total energy of excitation in direct proportion to the increase in the size of the stimulus. If it be conceived that the receptors act collectively one might expect an increase in the response of the sense organ in direct proportion to the increase in the area of the stimulus, provided that the receptors included in the area stimulated are equally sensitive. (*b*) If, on the other hand, it be conceived that the receptors act independently, it might be expected that the intensity per unit area of stimulus, not the total intensity, would be the effective factor in determining the response of the sense organ. In this case increase of area would not increase the power of the stimulus to arouse a response in the sense organ. (*c*) A third view would be that increase of area is effective in arousing the response of the sense organ, but that this effect is not in direct proportion to the increase of area. In a somewhat extensive investigation of the response of the retina made in a number of studies, we have found this latter view to be correct within limits. That is, within a certain range of change, an increase of size of stimulus increases the response of the sense organ, but apparently not in direct proportion to the change in size. Even if the first view as stated above correctly represents the law of response of nervous structures, the conditions for its fulfillment would not be found in the retina. That is, it presupposes a uniform distribution of sensitivity over the receiving surface. Uniform sensitivity occurs over very limited areas in the retina.

Next to intensity, size of stimulus is the most powerful of the extraneous factors influencing determinations of the extent and shape of the color fields. Three reasons may be mentioned why the perimetrist should be thoroughly familiar with this influence for all of the conditions under which it may occur. (*a*) Reproducibility of result cannot be expected without careful control of size of stimulus. (*b*) No result obtained with any given size of stimulus can be correctly interpreted without a knowledge of the effect of size of stimulus and of the range of variation that may be expected for non-pathological cases with that size of stimulus. And (*c*) the effect of size of stimulus is greater on the extent and shape of pathological than of non-pathological fields, i.e., the pathological field is narrowed more by the use of small stimuli than the non-pathological field. Advantage of this fact can be taken to increase the diagnostic sensitivity in the determination of the form and color fields. Other devices may also be used to increase diagnostic sensitivity. The discussion of this point, however, will not be carried further in this paper.

In one of our earlier papers on the effect of intensity of stimulus on the size and shape of the color fields we were able to show that the fields for red and blue can be made co-extensive with the limits of the field of achromatic vision if a sufficiently intensive stimulus be used. One of the points included in the present study has been to find out whether the same result can be obtained by increasing the area of stimuli of good but not of excessive intensity.

A detailed explanation of the effect of brightness of preexposure and surrounding field on the response of the peripheral retina to color has been given in previous papers (2, 6, pp. 110-157). Space will not be taken here, therefore, for a discussion of these effects.

Conditions Under Which the Work Was Done

The determinations were made with the Ferree-Rand perimeter (5). The illumination of the perimeter arm was kept constant at 7 f. c. Three brightness values were used as preexposure and surrounding field—a white, a black, and a gray of the brightness of the color at or near the point at which the limits of sensitivity occur. The brightness value of the white at the intensity of illumination employed was 0.01367 c. p. per sq. in.; of the black, 0.000567 c. p. per sq. in. The grays needed to match the colors in brightness were selected from the Hering series of pigment grays in which white is represented as Number 1 and black as Number 50. Inasmuch as the colors change their brightness when imaged at different degrees of excentricity from the fovea and the different sizes of stimulus carried the limits to different degrees of excentricity in the peripheral field, different grays were required to match the stimuli of different sizes at the limits of sensitivity. In some cases also these grays varied for the different meridians for the same color. The serial numbers of these grays, also their brightnesses in candle-power per square inch under the intensity of illumination used, are given in Table 1. The stimuli used were red, green, and blue of the Heidelberg series of pigment papers. The sizes employed subtended a visual angle at the eye of 2, 1, 0.5, and 0.17 deg., respectively.

As stated above, an attempt was made to find out whether the

TABLE 1

Showing the Serial Numbers and Brightnesses in c.p. per sq. in.* under 7 f.c. of Illumination of the Gray Pigments (Hering Series) Which Matched the Different Sizes of Stimuli at the Limits of Excentricity

Visual angle (deg.)	Red		Green		Blue	
	Serial number	C.p.per sq. in.	Serial number	C.p. per sq. in.	Serial number	C.p. per sq. in.
5	41	.00127	6	.00651	29	.00214
2	41	.00127	6 3	.00651 .00883	33 31	.00177 .00195
1	33	.00177	7	.00613	33	.00177
0.5	41	.00127	7	.00613	29	.00214
0.17	39	.00161	5	.00682	29	.00214

*Values specified in candle-power per square inch may be converted into millilamberts by multiplying by 486.8.

color fields could be made coextensive with the field of vision by increasing the area of stimuli of good but not excessive intensity. For this work only the gray preexposure and surrounding field were used. Various sizes of stimulus were investigated for each color. It was found that an increase in the visual angle subtended by the colored stimulus up to 5 deg. widened the limits of sensitivity to that color. An increase beyond this size widened the limits only in case the limit is defined as the degree of excentricity of the *center* of the stimulus and then only by the amount which is some approximation of half the visual angle subtended by the stimulus. The limits of sensitivity were therefore determined for this optimum size of stimulus, 5 deg., in order to show the widest fields that could be obtained for the intensity of stimulus employed and for the observer used by increasing the area of the stimulus. For comparison the field was also determined for a white stimulus subtending a visual angle of 1 deg. on a black background (coefficients of reflection, respectively, 78 and 4%).

In specifying the limits of sensitivity for stimuli of different area, two methods may be employed (6, pp. 6-12). (*a*) The usual method may be followed of specifying the limit as the degree of excentricity of the *center* of the stimulus when it is placed at the point at which all trace of its original quality is lost. Or (*b*) the limit may be defined as the degree of excentricity of the *inner edge* of the stimulus when it is placed at the point at which all

TABLE 2

SHOWING THE EFFECT OF SIZE OF STIMULUS ON THE AREA IN SQ. CM. OF THE COLOR FIELDS PLOTTED TO A SCALE OF 1 MM. TO 1° FOR RED, GREEN, AND BLUE STIMULI

Intensity of illumination, 7 f.c.

	Area in sq. cm. of color field for			
Color	2-deg. stimulus	1-deg. stimulus	0.5-deg. stimulus	0.17-deg. stimulus
		*Gray preexposure and surrounding field**		
Red	99.3	56.4	37.4	12.6
Green	36.0	16.8	11.4	6.5
Blue	94.2	54.6	40.25	12.5
		White preexposure and surrounding field		
Red	82.5	22.3	9.2	2.0
Green	33.3	13.1	7.5	2.5
Blue	85.6	27.0	17.9	2.5
		Black preexposure and surrounding field		
Red	92.5	48.8	30.0	8.2
Green	25.0	8.2	3.2	1.1
Blue	88.1	48.4	34.8	17.8

*Of the brightness of the color at the limits of sensitivity.

TABLE 3

SHOWING IN PERCENTAGE THE REDUCTION IN AREA OF THE COLOR FIELD WHEN STIMULI OF 1, 0.5, AND 0.17°, RESPECTIVELY, ARE USED, AS COMPARED, WITH THE AREA FOR A STIMULUS OF 2°

Intensity of Illumination, 7 f.c.

	Per cent reduction in area from the field for the 2-deg. stimulus for		
Color	1-deg. stimulus	0.5-deg. stimulus	0.17-deg. stimulus
	*Gray preexposure and surrounding field**		
Red	43.2	62.3	87.3
Green	53.3	68.3	81.9
Blue	42.0	57.3	86.7
	White preexposure and surrounding field		
Red	73.0	88.9	97.6
Green	60.7	77.5	92.5
Blue	68.5	79.1	97.1
	Black preexposure and surrounding field		
Red	47.2	67.6	91.1
Green	67.2	87.2	95.6
Blue	45.1	60.5	79.9

*Of the brightness of the color at the limits of sensitivity.

TABLE 4

Showing in Percentage the Reduction in Area of the Color Field When White and Black Preexposures and Surrounding Fields Are Used as Compared with Gray of the Brightness of the Color at the Limits of Sensitivity

Intensity of Illumination, 7 f.c.

Color	Percentage reduction in area from the field for gray preexposure and surrounding field for White preexposure and surrounding field	Black preexposure and surrounding field
	2-deg. stimulus	
Red	16.9	5.8
Green	7.5	30.6
Blue	9.1	6.5
	1-deg. stimulus	
Red	60.5	13.5
Green	22.0	51.2
Blue	50.6	11.4
	0.5-deg. stimulus	
Red	75.4	19.8
Green	34.2	71.9
Blue	55.3	13.5
	0.17-deg. stimulus	
Red	84.2	34.9
Green	61.5	83.1
Blue	86.2	....

TABLE 5

Showing for Comparison the Area in sq. cm. of the Color Fields Plotted to a Scale of 1 mm. to 1° for the Optimum Size of the Stimulus and of the Form Field (White on Black) for a 1° Stimulus

Preexposure and surrounding field of gray of the brightness of the color at the limits of sensitivity. Intensity of illumination, 7 f.c.

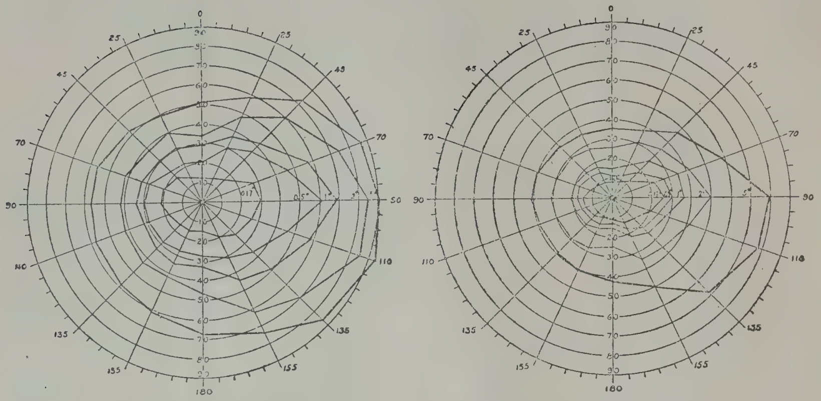
Stimulus	Size of stimulus (deg.)	Area of field (sq. cm.)
Red	5	156.5
Green	5	77.4
Blue	5	147.3
White	1	166.0

trace of its original quality is lost. The degrees of excentricity given by the two methods differ then by one-half the visual angle subtended by the stimulus. In this paper we have specified the limits of sensitivity in accordance with the first of the above-mentioned methods. The difference for the methods would be negligible for our results in case of the stimuli subtending visual angles of 0.17, 0.5, and 1 deg. For the stimuli subtending visual angles of 2 and 5 deg., the limits are wider by 1 and 2.5 deg. respectively, as we have specified them, than they would be by the second method.

RESULTS

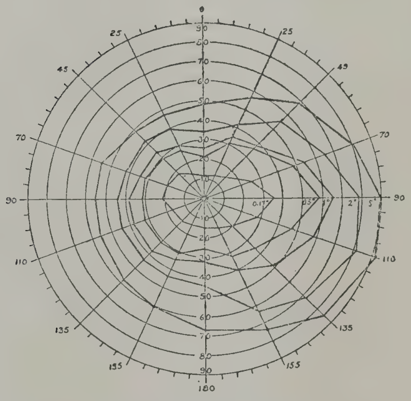
The results obtained are shown in Tables 2-5 and Figures 1-8. A precise way of making a quantitative comparison of fields for the sake of determining numerically the exact effect of any given factor is to measure the areas of the fields plotted to a standard scale and determine the relations of these areas. The measurements can be made most conveniently by means of a planimeter. The perimetrist will find this instrument of great value and convenience in compiling and interpreting his results. We have measured the areas of the fields shown in Figures 1-8 when plotted to a scale of 1 mm. to 1 deg. of visual angle, and have determined in percentage the amount the field has been changed (*a*) by each of the sizes of stimulus for the three conditions of preexposure and surrounding field, using the area of the field determined with the 2 deg. stimulus as the base on which to compute the percentage change; and (*b*) by each of the brightnesses of preexposure and surrounding field for the four sizes of stimulus, using as the base on which to compute the percentage change the area of the field for the preexposure and surrounding field of gray of the brightness of the color at the limits of sensitivity. In Table 2 are given the areas in sq. cm. of the fields when plotted to a scale of 1 mm. to 1 deg. for the different sizes of stimulus and for the different brightnesses of preexposure and surrounding field.

The effect of size of stimulus on the color fields is shown in Table 3 and Figures 1-4. In Table 3 is shown in percentages the reduction in area of the color fields for stimuli of 1 deg., 0.5 deg. and 0.17 deg., as compared with a stimulus of 2 deg. In Figure 1, Charts A, B, and C, fields are given for each of the three colors used for all sizes of stimulus with the gray preexposure and surrounding field; in Figure 2, Charts A, B, and C, with the white preexposure and surrounding field; in Figure 3, Charts A, B, C, with the black preexposure and surrounding field. In Figure 4, Charts A, B, C, and D, the fields are given for the three colors taken respectively with 2-deg., 1-deg., 0.5-deg., and 0.17-deg. stimuli and preexposure and surrounding field in each case of the gray of the brightness of the color at the limits of sensitivity.



A. Fields for Red

B. Fields for Green

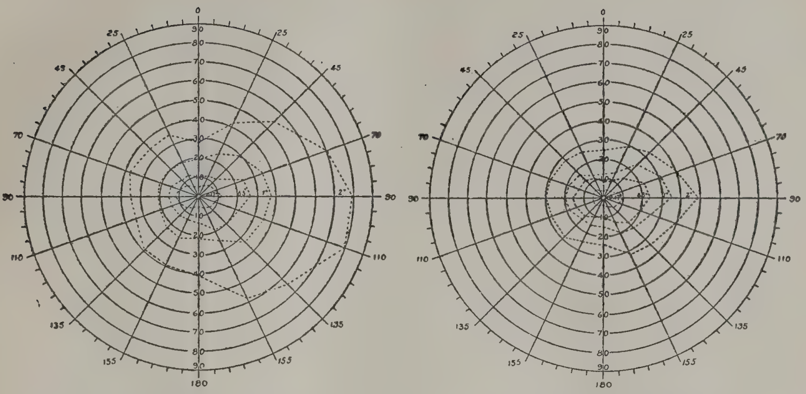


C. Fields for Blue

FIGURE 1

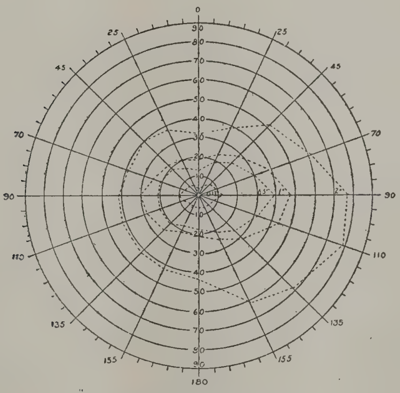
Showing the Effect of Size of Stimulus on the Extent and Shape of the Color Fields for Red, Green, and Blue

Preexposure and surrounding field of gray of brightness of color at limits of sensitivity. Intensity of illumination, 7 f.c.



A. Fields for Red

B. Fields for Green

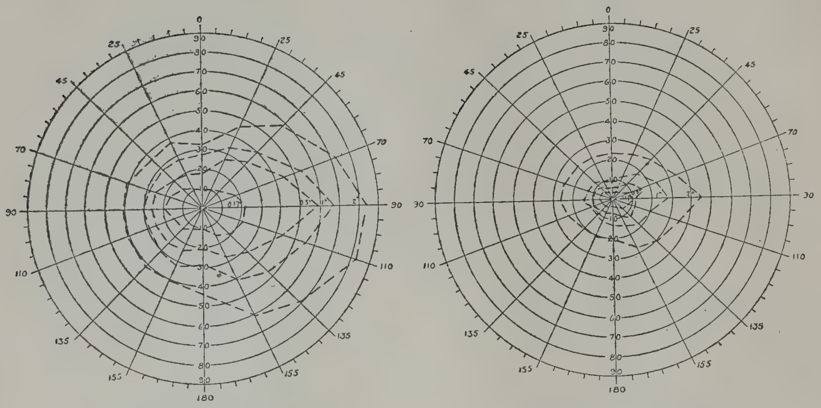


C. Fields for Blue

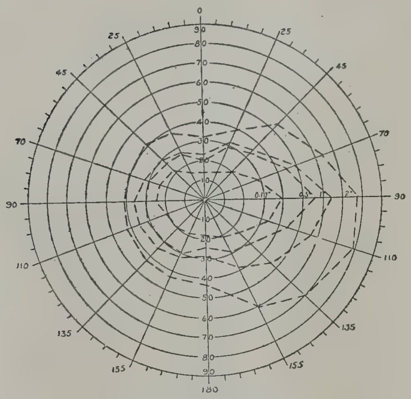
FIGURE 2

Showing the Effect of Size of Stimulus on the Extent and Shape of the Color Fields for Red, Green, and Blue

White preexposure and surrounding field. Intensity of illumination, 7 f.c.



A. Fields for Red B. Fields for Green

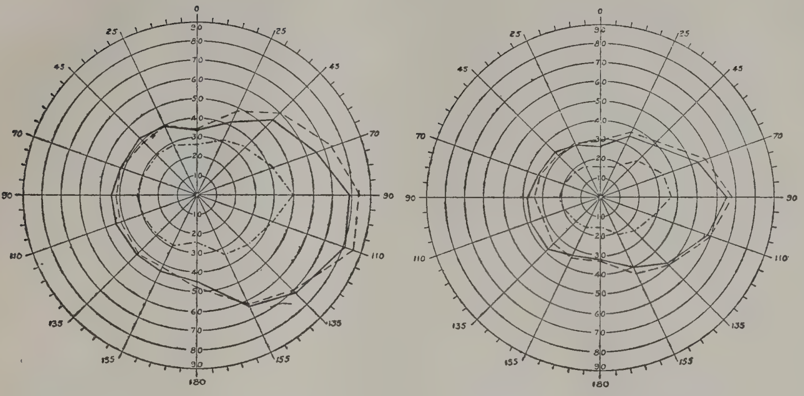


C. Fields for Blue

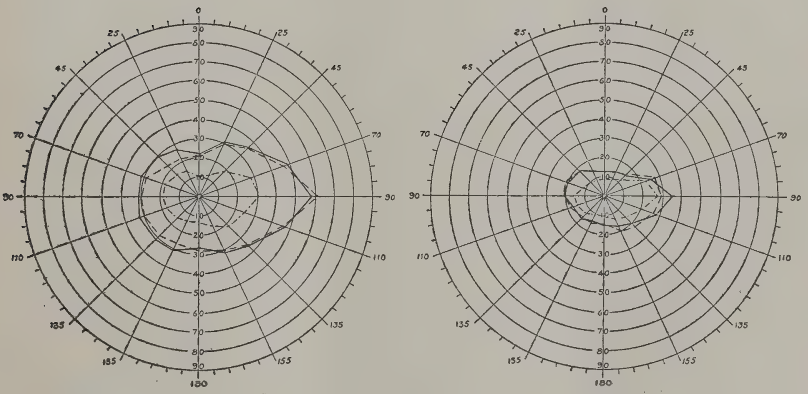
FIGURE 3

Showing the Effect of Size of Stimulus on the Extent and Shape of the Color Fields for Red, Green, and Blue

Black preexposure and surrounding field. Intensity of illumination, 7 f.c.



A. Fields for 2° Stimulus B. Fields for 1° Stimulus



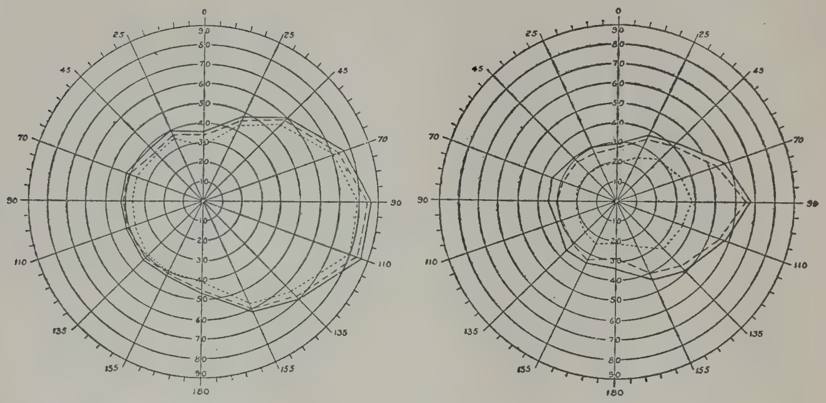
C. Fields for 0.5° Stimulus D. Fields for 0.17° Stimulus

FIGURE 4

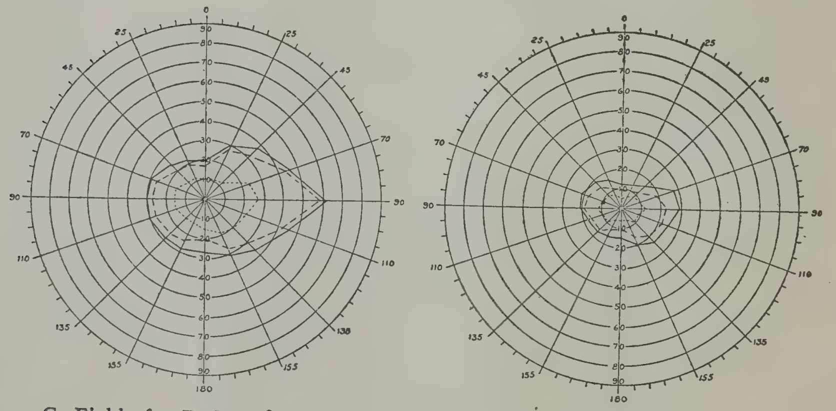
FIELDS FOR RED, GREEN, AND BLUE STIMULI SUBTENDING VISUAL ANGLES OF 2, 1, 0.5, AND 0.17°

Preexposure and surrounding field of gray of brightness of color at limits of sensitivity. Intensity of illumination, 7 f.c.

———— Blue, ------ Red, -·-·-·- Green



A. Fields for Red, 2° Stimulus B. Fields for Red, 1° Stimulus



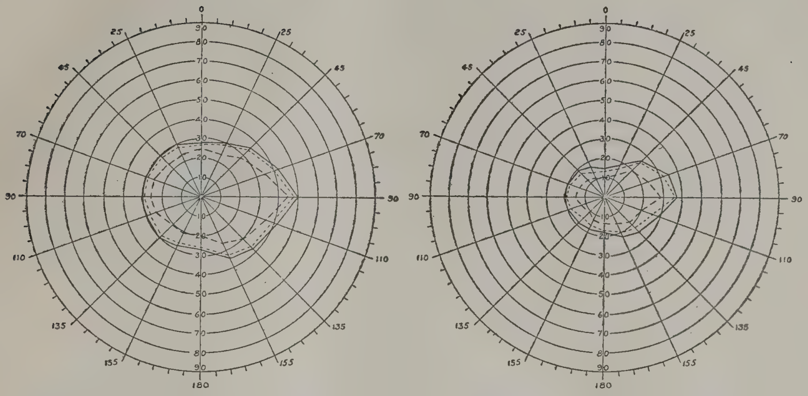
C. Fields for Red, 0.5° Stimulus D. Fields for Red, 0.17° Stimulus

FIGURE 5

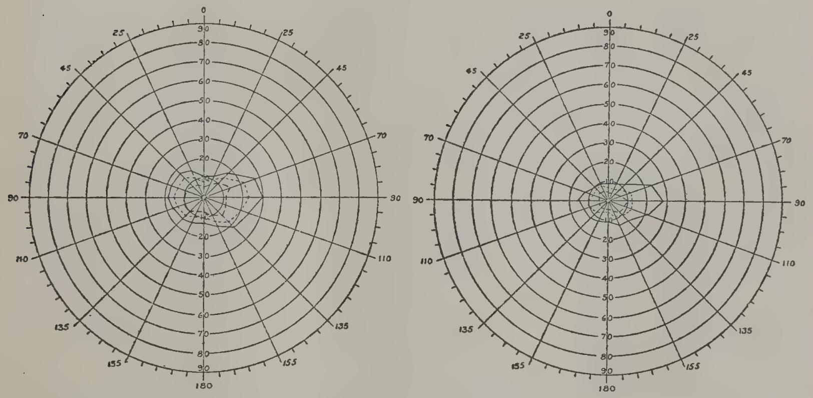
SHOWING THE EFFECT OF BRIGHTNESS OF PREEXPOSURE AND SURROUNDING FIELD ON THE EXTENT AND SHAPE OF THE COLOR FIELD FOR RED WITH STIMULI OF 2, 1, 0.5, AND 0.17°

Intensity of illumination, 7 f.c.

——— Gray, - - - - - - Black -·-·-·-· White



A. Fields for Green, 2° Stimulus B. Fields for Green, 1° Stimulus

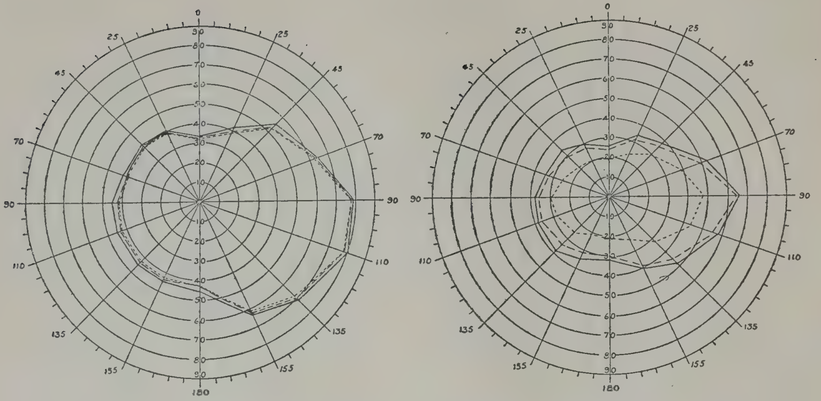


C. Fields for Green, 0.5° Stimulus D. Fields for Green, 0.17° Stimulus

FIGURE 6

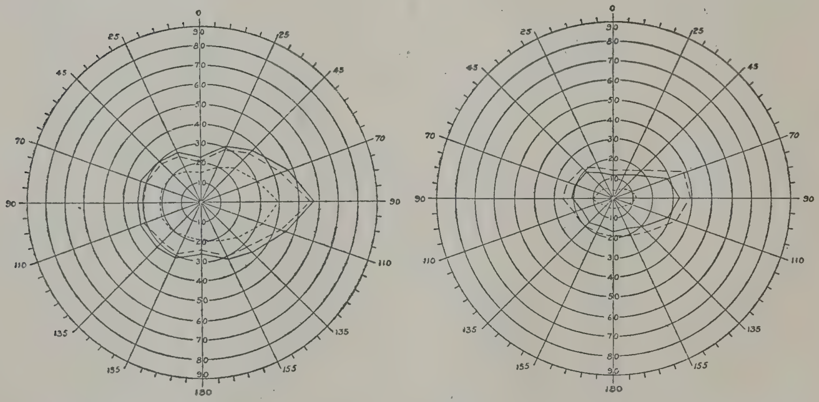
Showing the Effect of Brightness of Preexposure and Surrounding Field on the Extent and Shape of the Color Fields for Green with Stimuli of 2, 1, 0.5, and 0.17°

Intensity of illumination, 7 f.c.



A. Fields for Blue, 2° Stimulus

B. Fields for Blue, 1° Stimulus



C. Fields for Blue, 0.5° Stimulus

D. Fields for Blue, 0.17° Stimulus

FIGURE 7

Showing the Effect of Brightness of Preexposure and Surrounding Field on the Extent and Shape of the Color Fields for Blue with Stimuli of 2, 1, 0.5, and 0.17°

Intensity of Illumination, 7 f.c.

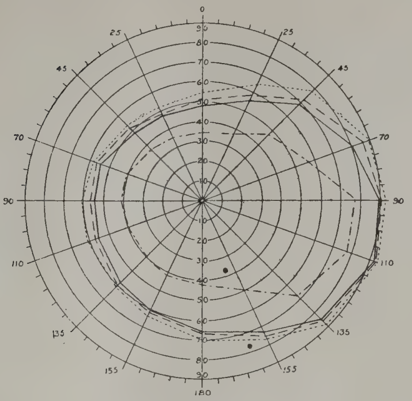


FIGURE 8

Showing the Fields for Red, Green, and Blue for Stimuli of Optimum Size (5°) and for the Form Field (White Stimulus on Black (1°)

Preexposure and surrounding field of gray of brightness of color at limits of sensitivity. Intensity of illumination, 7 f.c.

------ White, ———— Blue, — - — Red, — — — Green

The effect of the brightness of preexposure and surrounding field is shown in Table 4 and Figures 5-7. In Table 4 is shown in percentages the reduction in area of the color fields for white and black preexposures and surrounding fields as compared with a preexposure and surrounding field of a gray of the brightness of the color at the limits of sensitivity. In Figure 5, Charts A, B, C, and D, are given the fields for red with white, black, and gray preexposures and surrounding fields for each of the four sizes of stimulus used; in Figure 6, Charts A, B, C, and D, the fields for green; and in Figure 7, Charts A, B, C, and D, the fields for blue.

In Table 5 and Figure 8 the results for the optimum size of stimulus are given. In Table 5 a comparison is made of the area in sq. cm. of the fields for red, green, and blue with a 5-deg. stimulus (gray preexposure and surrounding field) and that of the form field (white on black) with a 1-deg. stimulus. In Figure 8 the fields are given for the three colors with a 5-deg. stimulus (gray preexposure and surrounding field) and for the form field (white on black) with a 1-deg. stimulus.

The results obtained may be briefly summarized as follows:

1) Size of stimulus is a powerful determinant of the visibility of color in the peripheral field, comparable with intensity in importance.

2) By increase of size of stimulus the limits of sensitivity to red and blue for stimuli of the intensity employed can be made almost co-extensive with the limits of the field of vision. With stimuli subtending a visual angle of 5 deg. illuminated by 7 f. c. of light, the fields for red were only 6 and for blue only 11% smaller than for a 1-deg. stimulus of white on black. Further increase in size of the colored stimulus for this intensity of illumination did not seem materially to increase the extent of the field. As has been demonstrated in former papers, the limits of these colors with higher intensities of stimulation are co-extensive with the limits of the field of vision for stimuli much smaller than 5 deg. It was one of the purposes of this paper to see whether this result could be accomplished by increasing the size of stimuli of moderate intensity.

3) Size of stimulus produced the least effect with a pre-exposure and surrounding field of the brightness of the color. The effect, as might be expected, was considerably greater for the white and black preexposures and surrounding fields. So far as we have been able to determine, the effect of any factor, non-pathological or pathological, is greater in the presence of other factors than when the influence of these factors is eliminated. This fact may contain a suggestion as to methods of increasing diagnostic sensitivity. Any use of a factor for that purpose, however, presupposes the standardization of its influence. Also, to be differential or of value for diagnosis, the effect must be greater on the pathological than on the non-pathological eye. This in general seems to be the case.

4) The effect of size of stimulus in the presence of brightness difference between stimulus and preexposure and surrounding field is shown in Table 2 to bear some unknown relation to the magnitude of that difference. That is, the effect in case of blue and red (the darker colors) was considerably greater for the white preexposure and surrounding field, and in case of green (the lightest color) for the black preexposure and surrounding field. A comparison of the coefficients of reflection of the grays which match these colors in brightness at the different positions in the field with those for the white and black used will show that the brightness difference between color and its preexposure and surrounding field was always greater in case of green with black and red and blue with white.

5) The widest angular limits of the color zones were obtained when the preexposure and surrounding field were of the same brightness as the color for all of the colors except blue with the very small

0.17-deg. stimulus. With this stimulus the widest field was obtained with the black preexposure and surrounding field. This result is quite exceptional. It seems to be obtained only when very small stimuli are used. Usually the widest fields are obtained with a preexposure and surrounding field of the brightness of color. Blue, however, sustains a very close qualitative relation to the gray of the Hering series which matches it in brightness. When very small in size it seems to escape observation entirely on such a background after a certain point of excentricity is reached. For this reason the small amount of induction from the black background seems to be a favorable rather than an unfavorable factor. That is, the influence of this induction on rendering the stimulus more different from its background seems to more than compensate for the small amount of inhibitive action on the chromatic component of the sensation.

6) With the white preexposure and surrounding field the size of the field for red with stimuli subtending visual angles of 2, 1, 0.5 and 0.17 deg. was reduced 16.9, 60.5, 75.4, and 84.2%, respectively; with the black preexposure and surrounding field, 5.8, 13.5, 19.8, and 34.9%. With the white preexposure and surrounding field for green these reductions were, respectively, 7.5, 22, 34.2, and 61.5%; with the black preexposure and surrounding field they were 30.6, 51.2, 71.9, and 81.3%. With the white preexposure and surrounding field the field for blue with stimuli subtending visual angles of 2, 1, and 0.5 deg., the reductions were 9.1, 50.6, and 55.3%; with the black preexposure and surrounding field they were 6.5, 11.4, and 13.5%. In the case of blue for the stimulus subtending a visual angle of 0.17 deg., as already stated, the field was widest for the black preexposure and surrounding field. The reduction produced by the gray preexposure and surrounding field was 29.6%; by the white preexposure and surrounding field, 86.2%.

7) It will be noted that much greater effects of preexposure and surrounding field were found for the small than for the large stimuli. In explanation of this the following points may be noted. (*a*) In general a decrease in the size of the stimulus adds to the sensitivity of the perimetric method or increases the ability to pick up the effect of any factor or condition, normal or pathological. For example, a given pathological condition causes a greater shrinkage of the field for small than for large stimuli. That is, a smaller amount of defect or disturbance can be picked up with a small than with a large stimulus. One reason for this is that the use of the

smaller stimulus results in a contraction of the field, thus bringing the limits into that region of the retina in which sensitivity falls off more gradually, and in which region, therefore, a given amount of effect on the color response produces a greater change in the field. Another reason is that a more difficult task is set for the eye by a small stimulus. This results in a diminished power of discrimination for a small as compared with a large stimulus under the unfavorable conditions imposed and therefore a greater effect of these conditions on the results of the test. (*b*) There is a more intensive achromatic induction over the smaller stimuli, i.e., the amount of induction is influenced by the ratio of the area of inducing surface to the area over which the induction occurs.

8) It will be noted that black as a preexposure and surrounding field contracts the limits more for the lightest color, green; and white for the darker colors, red and blue. The explanation for this difference in result is doubtless to be found in terms of the effect of the surrounding field. That is, the greater the brightness difference between the stimulus and its surrounding field, the greater will be the induction effect and the greater its inhibitive action on the color.

As a plea for the adoption of better controls in the practice of perimetry it may not be out of place to add in conclusion the following comments. In general, two types of method are available for determining the condition of a sense organ, the direct or objective examination and the functional test. Field study of every kind belongs to the latter type of examination. The following are some of the advantages of the functional test over any objective method of making the examination. (*a*) Perhaps the earliest manifestation of a pathological condition is a disturbance of function. This usually comes before an observable change in structure. (*b*) In case of the eye the visual field represents a projection of the retina. The areal magnification of this projection is in direct ratio of the square of the distance of the retina from the nodal point to the square of the distance of the nodal point to the projection field.[1] Through the application of the functional test in the study of the visual field, therefore, a much greater magnification can be secured than by any objective method of examining the nerve-fiber layer of the retina or fundus—that is, a much more minute and searching examination can be made. And (*c*), as just stated, the objective examination shows only the condition of the fundus. The condition of the deeper lying structures can not be revealed by it. The functional test furnishes evidence of the condition of the entire sensorium from retina to brain center. (*d*) Results by the subjective methods of examination can be subjected more or less satisfactorily to numerical treatment and comparison. And (*e*) the subjective method of examination deals with a static, the functional test with a dynamic condition—a dynamic condition of the entire sensorium. The functional tests in proportion to their comprehensiveness and sensitivity scrutinize minutely the sensorium in action and there is little in the way of depression or disturbance that should escape their survey.

---

[1] For a distance of 33 cm. this magnification amounts approximately to 484; for a distance of 1 m., 4444; and for a distance of 2 m., 17,778.

Obviously the minuteness of the subjective examination is limited only by the threshold of perception. In the central field an observer with an acuity of 20/20 can discriminate a detail of image of a breadth of approximately 0.00001 mm.; with an acuity of 30/20 a detail of the breadth of approximately 0.0000067 mm.; and with an acuity of 40/20 a breadth of detail of approximately 0.000005 mm.

It is obvious, therefore, that there is great need to develop the possibilities of the functional test as a means of differential study and diagnosis. A sensitive and precise means of early diagnosis is, it scarcely need be pointed out, one of the most important adjuncts of preventive medicine.

References

1. Ferree, C. E., & Rand, G. The absolute limits of color sensitivity and the effect of intensity on the apparent limits. *Psychol. Rev.,* 1920, **27**, 1-23.
2. ——————. The limits of color sensitivity; effect of brightness of preexposure and surrounding field. *Psychol. Rev.,* 1920, **27**, 377-398.
3. ——————. Factors which influence the color sensitivity of the peripheral retina. *Trans. Amer. Ophth. Soc.,* 1920, **18**, 171-197.
4. ——————. The effect of variation of the intensity of the illumination of the perimeter arm on the determination of the color fields. *Psychol. Rev.,* 1922, **29**, 457-473.
5. ——————. A new laboratory and clinic perimeter. *J. Exper. Psychol.,* 1922, **5**, 46-67.
6. Rand, G. The factors that influence the sensitivity of the retina to color: a quantitative study and methods of standardizing. *Psychol. Monog.,* 1913, **15**. Pp. xi+169.

The Johns Hopkins University
Baltimore, Maryland

LA GRANDEUR DU STIMULUS ET LA CLARTÉ DE L'EXPOSITION D'AVANCE ET DU CHAMP ENVIRONNANT DANS LEURS RAPPORTS AVEC L'ÉTENDUE ET LA FORME DES CHAMPS DE COULEUR

(Résumé)

La méthode ordinaire et approuvée de déterminer l'influence d'un facteur variable est de tenir constants à quelque valeur déterminée d'avance tous les facteurs excepté celui dont on désire étudier les effets. On a une idée très incomplète de l'effet si l'enquête reste à ce point. C'est-à-dire, on obtient des résultats très différents quand on choisit des valeurs différentes pour un ou plus des facteurs qu'on tiendra constants pendant la série. Cet article montre l'effet de la grandeur du stimulus sur l'étendue et la forme des champs de couleur; aussi l'action intermédiaire de l'effet de la grandeur du stimulus et de la clarté de l'exposition d'avance et du champ environnant. On a fait des déterminations au moyen du périmètre Ferree-Rand avec une illumination de 7 pieds.—c. Les résultats montrent que la grandeur du stimulus est un déterminant puissant de la visibilité de couleur dans le champ périphérique, comparable à l'intensité en importance; que, si l'on augmente la grandeur jusqu'à un angle visuel de 5 degrés, les limites de la sensibilité au rouge et au bleu de l'intensité employée peuvent avoir presque la même étendue que les limites de la vision de lumière blanche; que la grandeur produit le moindre effet avec l'exposition d'avance et le champ environnant de la couleur; que l'effet de la grandeur en présence d'une différence de clarté entre le stimulus et l'exposition d'avance et le champ environnant a un rapport très important avec la grandeur de cette différence; et que les petits stimuli ont montré de plus grands effets de l'exposition d'avance et du champ environnant que les grands stimuli.

FERREE ET RAND

GRÖSSE DER REIZE UND HELLIGKEIT DER VORBELICHTUNG UND DES UMGEBENDEN FELDES IN BEZUG AUF DIE AUSDEHNUNG UND DIE FORM DES FARBENFELDES

(Referat)

Beim Bestimmen des Einflusses eines variablen Faktoren ist das gewöhnliche und anerkannte Verfahren das, alle Faktoren mit Ausnahme desjenigen, dessen Effekte der Untersuchung unterliegen auf einem vorausbestimmten Wert konstant zu erhalten. Man hat aber ein sehr unvollständiges Bild der Effekte, wenn die Untersuchung damit aufhört. D. h. man erhält sehr verschiedene Ergebnisse, wenn man verschiedene Werte wählt für einen oder mehrere Faktoren, die während der Serie konstant zu halten sind. Diese Arbeit zeigt den Effekt der Reizgrösse bezüglich Ausdehnung und Form der Farbenfelder; ferner die Verwebung der Effekte aus Grösse der Reize und Helligkeit der Vorbelichtung und umgebenden Feld. Man erhielt die Bestimmungen auf dem Ferree- Rand- Perimeter, mit einer Beleuchtung von 7 Fusskerzen. Die Ergebnisse zeigen, dass Reizgrösse ein sehr wirksamer Bestimmungsfaktor ist bezüglich der Sichtbarkeit von Farben im peripheren Feld, mit Hinschit auf die Wichtigkeit der Intensität vergleichbar; ferner dass bei wachsender Grösse des Seh-

winkels bis zu 5 Grad die Grenzen der Empfindlichkeit bezüglich rot und blau bei der angewandten Intensität fast auf die Grenzen des Weisslichtsehens gebracht werden können; und dass die Grösse den geringsten Effekt hervorruft wenn die Vorbelichtung und das umgebende Feld von derselben Helligkeit sind wie die Farbe; und dass der Effekt der Grösse in der Gegenwart von Helligkeitsdifferenzen zwischen Reiz und Vorbelichtung und umgebendem Feld eine wichtige Beziehung hat in Bezug auf jene Differenzen; endlich dass für kleine Reize grössere Effekte der Vorbelichtungen und des umgebenden Feldes gefunden wurden als für grosse Reize.

Ferree und Rand

EFFECTS OF VISUAL EXPOSURE ON THE RATE AND RELIABILITY OF STYLUS-MAZE LEARNING*[1]

From the Jesup Psychological Laboratory, George Peabody College for Teachers

Joseph Peterson and Loy W. Allison

In view of recent discussions of the reliability of different learning processes by animals in mazes and also by humans in the stylus maze [Carr (2), Heron (3, 4), Hunter (5, 6), Hunter and Randolph (7), Paterson (10), Stone and Nyswander (12), Tolman and Nyswander (13), and Tryon (14)], it seems desirable to attempt improvements in the reliability of learning processes generally, insofar as this is possible without decreasing certain scientific values of these processes. Stone and Nyswander have obtained high reliability coefficients in multiple-T mazes for rats by preventing return runs. Nyswander (9) has found that reliability is probably increased in the high relief finger maze for humans as compared with the stylus maze, if degree of difficulty be equalized. Our work is rather closely related to that of Carr (1) on the influence of visual guidance in maze learning, except that we are stressing particularly the question of improving the reliability. He found that errors were markedly reduced in the sort of stylus maze used (invisible stops) by him and that more trials were required for learning when one minute of visual exposure of the maze was allowed just *before* the tracings; that visual aid *during* the first and the second tracings reduced errors by over 50%, but did not reduce trials; and that vision of the maze *during* Trials 3 and 5 reduced the number of trials 40 and of errors 80%, as compared with the control group which was allowed no visual exposure.

A reduced form of the maze used in this experiment is shown in Figure 1. The maze, 9 x 12 inches and one-fourth of an inch thick, was made of a rubberized composition. The grooves, one-fourth of

---

*Received for publication by Carl Murchison of the Editorial Board, November 12, 1929.

[1] Mr. Allison took all the subjects through the tracings of the maze under the direction of the senior author. The authors collaborated in the working up of the data. The senior author is responsible for this statement of results.

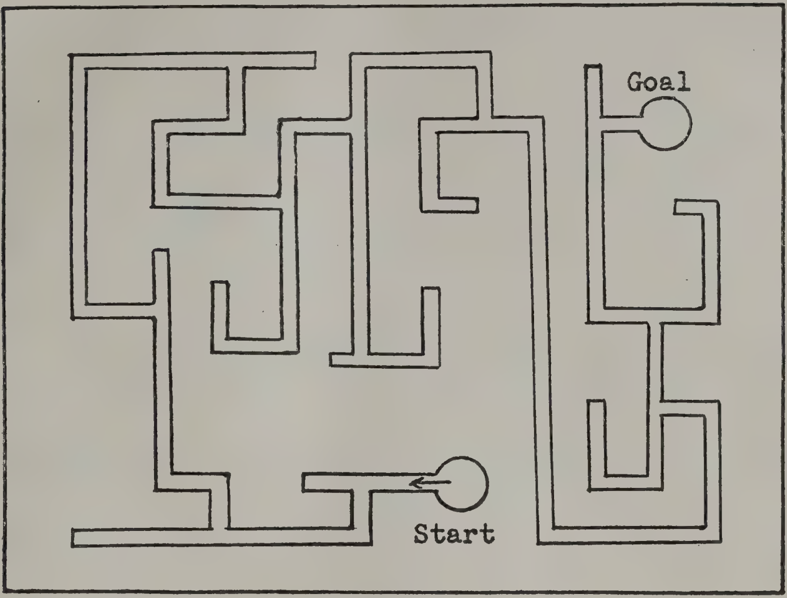


FIGURE 1
STYLUS MAZE A

an inch wide, were cut through the composition. The maze was placed on a piece of heavy plate-glass which formed a smooth surface for the stylus to move upon, and was securely fastened to the top of a table.

S was seated directly in front of the table, in a straight-back chair, the middle of her body being in line with the middle of the maze. No effort was made to keep constant the distance between *S* and the table, this distance being determined by the length of *S*'s arm. *S* was merely instructed to "Sit in a comfortable position as for writing."

No oral instructions were given *S* at any time during the experiment, but a typewritten sheet of instruction, giving a general description of the procedure, was handed each *S* before the experiment began. Four groups of *S*'s, Groups 0, 1, 2, and 3, to be described later, were used.

The sole instructions for Group 0 were as follows:

> "A maze, as you know, is a winding pathway which leads to a goal. If

you depart from the true pathway, however, you run into blind alleys. The object of this experiment is for you to learn to trace your way with the stylus from the entrance of the maze to the goal without entering a single blind alley.

"You will be blindfolded, and therefore will not be able to see your way. You must simply explore in whatever fashion you think best until you find your way to the goal.

"I shall set your hand with the stylus at the entrance to the maze, and give you the signal to start. You must never raise your stylus from the base of the maze, nor jump over the sides of the pathway. Grip the stylus easily and hold it in a vertical position as far as possible.

"Don't become discouraged if you make numerous errors at first; these are inevitable. Don't stop trying until I say 'Goal,' when I shall raise your hand from the maze and give you a brief rest before the next trial.

"Ask no questions about the maze. Sit in a comfortable position as for writing. Be sure your fingers never touch the maze."

In addition to the above instructions, which were given alike to all *S*'s of each group, the following were included for Groups 1, 2, and 3, except for the differences noted presently.

"Before you begin each trial you will be allowed to study the maze for five seconds. During this time, however, you are to make no motion with your hands either over the maze or as if tracing an imaginary maze."

Where "5 seconds" occurs in the above statement for Group 1 "10 seconds" and "20 seconds" were inserted for Groups 2 and 3, respectively.

A complete record was kept of the tracings of each individual in each group. This record consisted of the time in seconds taken for each trial and the number and place of blind-alley entrances during each trial. During the progress of the tracing the experimenter was seated at the side of the table where each move of *S* could be carefully noted. Each entrance of the stylus into a blind-alley by even as much as one-fourth of an inch was called an error. To check *E*'s accuracy in recording errors one subject not in any of the four groups was used and required to learn the maze under blindfold conditions. *E* not only kept a record as he had done with the other *S*'s, but a clean sheet of paper was also placed under the maze before each trial and a pencil was used for tracing, instead of the stylus. This gave an objective self-record of each trial. The results were tabulated and the errors at different blind alleys were correlated with the results of the record kept by *E*. This gave a correlation of .99±.03. Only one difference was found between the two records, and even this difference may have been due to the inability at places to count the pencil marks easily, especially when there was much retracing. No further test of this point was made as it seemed evident that *E*'s counting was highly reliable.

A large rectangular piece of cardboard, 16½ x 19 inches, was used to cover the maze when not in use. And with Groups 1, 2, and 3, this was used as a cover board, which was raised to a vertical position on the side of the maze opposite the subject and held in this position by *E* during the period of visual inspection by *S*. The cover was quickly lowered at the end of the period of inspection, and the blindfold again replaced.

A piece of dark cloth four inches wide and having five folds was used as a blindfold. This was carefully placed over *S*'s eyes and fastened behind the head in such a manner as to make any vision impossible.

Sixty *S*'s were used, all students of George Peabody College. Of this number 50 were from the elementary psychology class, 2 were from a class in experimental psychology, and 8 were from a class in elementary education. All of the *S*'s were women, the purpose of this use of only one sex being to avoid any possible sex differences which might exist. None of the subjects had any previous experience with the maze problem in any way. Three of the groups, for whom we have intelligence test scores (Groups 0, 2, and 3), were of equal ability in the Alpha test (see Table 5).

The four groups of subjects each consisted of 15 students. These groups will be referred to as Groups 0, 1, 2, and 3. Group 0 learned the maze under strictly blindfold conditions. The members of Group 1 were each given 5 seconds, those of Group 2, 10 seconds, and those of Group 3, 20 seconds to study the maze before the first trial and before each succeeding trial until the maze was learned, as shown in the instructions.

The coefficients of reliability obtained in this experiment are given in Table 1. Three methods were used in their determination: the

TABLE 1

RELIABILITY COEFFICIENTS OBTAINED IN THREE WAYS ($N = 15$)

Group	Seconds exposure	Odds-Evens Rank differences	Odds-Evens Product moment	Av. of individual *r*'s
0	0	.42±.15	.48±.13	.68
1	5	.89±.04	.75±.08	.59
2	10	.96±.01	.87±.04	.66
3	20	.81±.06	.87±.04	.75

odds-evens (*a*) rank-difference and (*b*) product moment methods, and (*c*) the average of the individual correlations by the rank difference method. In the first two of these the total errors by all the *S*'s on the odd trials were correlated with the total errors on the even trials. In this case it was desirable not to include records after the point of zero errors had been reached, so only the first 16 trials were included in the count, $r_{(1+3+5....+15)\ (2+4+6....+16)}$. In the third method the correlations given are the averages (for the different periods of exposure) of all the correlations of errors in odd with those in even trials of the several *S*'s. This average correlation, it will be noted, is considerably larger for Group 0, and smaller for Groups 1, 2, and 3, (5, 10, and 20 seconds visual exposure, respectively), than are the corresponding total-error correlations by the usual methods. The distributions of the individual correlations are shown in Table 2, where the tendency for the correlations of the experimental groups to rise with length of exposure up to 20 seconds is evident. The first two methods show a marked and reliable gain in reliability, even for the small number of cases in each group ($N=15$), in the cases of the groups who were allowed to see the maze for just a few seconds before each trial (experimental groups). The gain is marked for even 5 seconds exposure, and increases only a little more when 10 seconds are allowed. Beyond this, probably because of the greater ease in learning with longer exposures, there

TABLE 2

Distribution of Individual Correlations: Odds-Evens Method

	Groups			
r	0	1	2	3
1.00		1	1	3
.90	2	1	2	3
.80	4	3	3	3
.70	2	2	4	2
.60	1	1	1	1
.40	1	1		1
.30		1		
.20	1	1	1	1
.10		1		
.00			1	1
.50	4	2	1	
—.10		1	1	
Total	15	15	15	15

is no gain, and even a slight loss, in reliability with increase of exposure time. (These conditions are reversed when the method of averaging the individual correlations of errors in odd and even trials is used. Why this is the case we cannot say. By this method the number of trials included in the calculations varied for different individuals, extending in each case to the point of including one errorless trial. This method is unnecessary where only group comparisons are made, but for individual comparisons and predictions of individual efficiency it is essential. The constancy of individual ranks cannot be revealed by the gross methods of merely summing errors in different trials or at different blind alleys. The relations of averaging the individual correlations, on the one hand, and of finding correlations by the mere summary of errors by all *S*'s, on the other, needs careful experimental analysis.)

Our results suggest the possibility of increasing the reliability of human stylus mazes (and why not of high relief mazes as well?) by suitably chosen visual exposure periods in advance of each trial, the length of these varying, of course, with the difficulty of the maze pattern used. In the standardization of these mazes this factor merits careful consideration among other factors of which account must be taken. We did not have the T-mazes available at the time of the experiment, but our findings, when stated in this general way, will doubtless hold for this improved pattern also. The value of the stylus and finger mazes as means of comparing individuals and groups can probably be considerably increased by a standardization of these mazes which takes account of optimum exposure periods before each trial.

Table 3 summarizes the effects of visual exposure on learning rate in terms of number of trials, of errors, and of seconds to complete the learning to the point of two errorless tracings in three trials. The table shows, in the case of each of these criteria, (*a*) the mean number, (*b*) the coefficient of variability *M. V./M,* (c) the size of the difference between the mean of Group 0 and that of each of the other groups (M_x in the table) in terms of the *M. V.* of Group 0 (therefore putting this size into a form that is comparable whatever be the absolute size of the several criteria units), and (*d*) the critical ratio, the difference mentioned in (*c*) divided by its *P. E.*

As to trials, variability is decreased in the case of the 5-seconds exposure but reaches its largest value, (i.e., the test reaches its greatest

TABLE 3

EFFECTS OF VISUAL EXPOSURE ON RATE OF LEARNING AS JUDGED BY NUMBER OF TRIALS, OF ERRORS, AND OF SECONDS

Group	Seconds exposure	Mean	Trials			Mean	Errors			Mean	Seconds		
			$\frac{M.V.}{M}$	$\frac{M_o-M_x}{M.V._o}$	$\frac{M_o-M_x}{P.E._d}$		$\frac{M.V.}{M}$	$\frac{M_o-M_x}{M.V._o}$	$\frac{M_o-M_x}{P.E._d}$		$\frac{M.V.}{M}$	$\frac{M_o-M_x}{M.V._o}$	$\frac{M_o-M_x}{P.E._d}$
0	0	47	.49			276	.47			2280	.33		
1	5	27	.26	.87	3.81	136	.45	1.08	4.45	1455*	.35	1.09	4.15
2	10	26	.58	.91	3.75	129	.68	1.13	4.27	1162*	.45	1.47	5.53
3	20	19	.42	1.22	5.34	88	.57	1.28	6.18	1095*	.44	1.56	6.01

Note: in this table M_o-M_x = mean of Group 0 *minus* mean of Group 1, of Group 2, or of Group 3, as the case may be; $M.V._o$ is the mean variation of Group 0; and $P.E._d$ is the probable error of the difference of the means of the two groups in consideration.

*These means become 1590, 1422, and 1475 for Groups 1, 2, and 3, respectively, when total exposure time is included. The difference between these means and 2280 are still reliable.

power to differentiate the *S*'s) in the 10-second group. The same situation is found both for errors and for time, with the one exception that variability does not fall off in the latter case even in the 5-second group. From this standpoint, apparently, the 10-second exposure period is the best of the three tried. The next column, $\frac{M_o - M_x}{M.V._o}$ under each criterion, shows, as would be expected, a constant increase in ease of learning with increase in exposure time. The reduction in each of the three experimental groups is greatest in the case of time, next in errors, and least in trials, when it is measured in terms of the *M.V.* of the control group; but when this reduction is stated in terms of the percentage of decrease, in each of these criteria, over the mean of the control group we get the results shown in Table 4. By this latter method all three of the experimental groups show the greatest saving in errors; two of them show the next greatest saving in trials; and one of them (Group 2), in time. We cannot here profitably, without further data, enter into a discussion of the relative merits of these two methods of measuring the amount of saving in the several criteria (trials, errors, and time), since that would involve us in topics still to be worked out experimentally [e.g., the correlations between means and variabilities (cf. Peterson and Barlow, 11), as well as absolute zero points]. The saving in errors would probably be about as great as in any other criterion, since by one method it is greatest and by the other it is next greatest.

TABLE 4

PERCENTAGE OF SAVING IN THE SEVERAL LEARNING CRITERIA WITH DIFFERENT EXPOSURE PERIODS

Seconds exposure	Trials	Errors	Time
5	42.6	50.7	36.2
10	44.8	54.3	49.0
20	59.6	68.1	51.5

The decrease in the means of the three experimental groups over the mean of the control group (no visual exposure) is reliable in every case so far as the errors and time criteria are concerned and they would probably also be reliable in the case of trials if the number of cases in each group had been larger. At any rate, there is evidence of a marked saving in the experimenter's time when the visual exposure method (whatever its optimum length should be in any case) is introduced. This, with the increase in reliability by this method, justifies its use. With the same degree of difficulty in the mazes by the two methods, the reliability would doubtless be increased a good deal more than our results show when visual exposure periods are allowed, even if the time of exposure is included as the time of learning. Nyswander (9) found that while in some cases a high relief finger maze may give more reliable results than a stylus maze, this effect of change in kind of maze of the same pattern is not constant. She suggests that even though the high relief maze permits *S* to note *all* bifurcations in the path followed while the stylus maze can be followed only by contact with one side at a time, this superiority of the high relief finger maze does not always assure greater reliability, because, as in the present case, such other factors as ease of learning must also be considered. With proper control of the degree of difficulty of the two kinds of mazes, the high relief finger maze with an optimum of visual exposure before the first or preceding every trial (as the case might require) would doubtless prove to be the more reliable measure of the subjects' abilities. With the finger maze thus standardized for humans it would doubtless be a satisfactory piece of apparatus for work on this sort of learning.

TABLE 5

INTERCORRELATIONS OF THE SEVERAL CRITERIA IN STYLUS MAZE LEARNING AND CORRELATIONS OF THESE CRITERIA COMBINED WITH SCORES ON ARMY ALPHA

		Visual expo-sure	Intercorrelation of the three criteria				Correlations of equally weighted composite maze	Army	Alpha
Group	*N*	(sec.)	Trials-errors	Trials-time	Errors-time	Mean	score and Alpha	Mean	*M.V.*
0	15	0	.89	.76	.86	.84	.17	137.6	20.29
1	15	5	.72	.46	.58	.59	no Alpha data		
2	15	10	.64	.70	.79	.71	.23	137.7	19.67
3	15	20	.74	.60	.85	.73	.02	138.0	14.93

Table 5 shows that intercorrelations of the three criteria used in the present experiment are slightly reduced by visual exposure. This itself may be a valuable thing as they will then not all so nearly measure the same traits. Alpha scores of three of our groups were available, and show a practical equality of these three groups in intelligence as thus determined. Group 3, however, shows considerably less variability than the other two groups for which we have scores. Intelligence tests scores were correlated with a combined equally weighted score in the three criteria of the maze learning with the results shown in the table. Number of trials, of total errors made, and of total seconds required by the several *S*'s in each group were ranked inversely according to size (since a small score is a good score in these criteria). These ranks were added for each *S* and she was again ranked on the size of these average ranks. The final rank was then correlated with the intelligence scores. It is unfortunate that we did not have test scores for Group 1. The results from the data available show generally low but positive correlations of efficiency in maze learning with efficiency in intelligence test scores. While the correlations are not reliable, there is some indication that they might be raised by optimum visual exposures. How much the high relief finger T-maze with optimum visual exposures will contribute toward a further increase in these correlations of maze learning with intelligence, as evaluated by some good criterion, is a problem worth attacking. It is desirable that we have many tests with low intercorrelations each of which yields high correlations with intelligence.

SUMMARY

The present study, continuing recent investigations of the reliability of learning scores in mazes, employs the stylus maze on 60 college women, divided into four groups of 15 each, a control group (0) and three experimental groups (1, 2, and 3) all of equal ability on the Alpha test. The subjects all learned when blindfolded to trace a complex stylus maze under slightly different conditions for each group. Group 0 had no direct visual aid whatever, while Groups 1, 2, and 3 were allowed 5 seconds, 10 seconds, and 20 seconds, respectively, of visual exposure just before each trial. Criteria of maze ability used are number of trials, of errors, and of seconds to reach the point of two errorless tracings in three trials. The experimental groups, by the method of correlating total errors in odd with total errors in even trials, show marked increases in reliability over the control group (.75, .87, and .87, for 5-, 10-, and 20-second exposures, respectively, as compared with .48). Visual exposure also resulted in marked reductions in the difficulty of learning. If the degree of difficulty in learning under the conditions

allowed each group had been equated, the gain in reliability of scores by visual aid would doubtless have been even greater. The authors suggest that a complex high relief T-maze standardized for optimum visual exposure periods would doubtless make a very satisfactory maze-learning experiment for humans, probably affording fair correlations with some good measure of intelligence. Various tests of low intercorrelations, each of which has high or fair correlation with intelligence, by some good criterion of the latter, are needed.

REFERENCES

1. CARR, H. A. The influence of visual guidance in maze learning. *J. Exper. Psychol.*, 1921, **4**, 399-417.
2. ————. The reliability of the maze experiment. *J. Comp. Psychol.*, 1926, **6**, 85-93.
3. HERON, W. T. The reliability of the inclined plane problem box as a method of measuring the learning ability of the rat. *Comp. Psychol. Monog.*, 1922, **1**, No. 1, Part 1. Pp. 1-36.
4. ————. Individual differences in ability versus chance in the learning of the stylus maze. *Comp. Psychol. Monog.*, 1924, **2**, No. 7. Pp. 60.
5. HUNTER, W. S. Correlation studies with the maze in rats and humans. *Comp. Psychol. Monog.*, 1922, **1**, No. 1. Part 2. Pp. 37-56.
6. ————. A note in answer to Professor Carr's criticism. *J. Comp. Psychol.*, 1926, **6**, 393-398.
7. HUNTER, W. S., & RANDOLPH, V. Further studies in the reliability of the maze with rats and humans. *J. Comp. Psychol.*, 1924, **4**, 431-442.
8. MILES, W. The high relief finger maze for human learning. *J. Gen. Psychol.*, 1928, **1**, 3-14.
9. NYSWANDER, D. B. A comparison of the high relief finger maze and the stylus maze. *J. Gen. Psychol.*, 1929, **2**, 273-288.
10. PATERSON, D. G. The Johns Hopkins circular maze studies. *Psychol. Bull.*, 1917, **14**, 294-297.
11. PETERSON, J., & BARLOW, M. C. The effects of practice on individual differences. *27th Yrbk. Nat. Soc. Stud. Educ.*, Part II, 1928, 211-230.
12. STONE, C. P., & NYSWANDER, D. B. The reliability of rat learning scores from the multiple-T maze as determined by four different methods. *Ped. Sem.*, 1927, **34**, 497-524.
13. TOLMAN, E. C., & NYSWANDER, D. B. The reliability and validity of maze measurements for rats. *J. Comp. Psychol.*, 1927, **7**, 425-460.
14. TRYON, R. C. Demonstration of the effect of unreliability of measurements on a difference between groups. *J. Comp. Psychol.*, 1928, **8**, 1-22.

George Peabody College for Teachers
Nashville, Tennessee

LES EFFETS DE L'EXPOSITION VISUELLE SUR LA VITESSE ET LA CONSTANCE DE L'APPRENTISSAGE DU LABYRINTHE À STYLET

(Résumé)

Cette étude fait des recherches sur la vitesse et la constance de l'apprentissage du labyrinthe à stylet avec différentes quantités d'aide visuelle. Les sujets ont été soixante étudiantes universitaires, divisées en quatre groupes, chacun de quinze—un groupe de contrôle (0) et trois groupes expérimentaux (1, 2, et 3) égaux en intelligence dans le test Alpha. Tous les sujets ont eu les yeux bandés pendant le tracement; les membres du Groupe 0 n'ont pas eu d'aide visuelle, tandis que l'on a permis aux groupes 1, 2, et 3 de voir le labyrinthe pendant 5, 10, 15 secondes respectivement directement avant chaque épreuve. Comme critères de capacité à apprendre, on a employé le nombre d'épreuves, d'erreurs, et de secondes exigées à arriver au point de deux tracements sans erreur en trois épreuves. Le degré de constance des résultats est devenu beaucoup plus grand dans les groupes expérimentaux (0,75, 0,87 et 0,87 pour une exposition de 5, 10 et 15 secondes respectivement, comme comparé à 0,48 pour le groupe de contrôle). La méthode a été la corrélation des erreurs en les épreuves paires et impaires. L'exposition visuelle a donné aussi une grande réduction de la difficulté de l'apprentissage. Si l'on avait rendu égal cet facteur de difficulté, le gain de constance par l'aide visuelle aurait été sans doute plus grand. Les auteurs suggèrent qu'un labyrinthe complexe T de haut relief, standardisé pour les périodes de meilleure exposition visuelle, ferait une expérience satisfaisante de l'apprentissage du labyrinthe par des êtres humains, donnant probablement d'assez bonnes corrélations avec quelque bonne mesure d'intelligence. Pour l'usage général, il serait désirable d'avoir plusieurs tests de différentes espèces lesquels ont des corrélations peu élevées entre soi mais dont chacun donne une corrélation élevée avec quelque bonne mesure d'intelligence.

PETERSON ET ALLISON

DIE WIRKUNGEN DER VISUELLEN EXPOSITION AUF DEN GRAD UND DIE ZUVERLÄSSIGKEIT DER STYLUS-LABYRINT-GEWÖHNUNG

(Referat)

Diese Untersuchung bezieht sich auf den Grad und die Zuverlässigkeit der Stylus-Labyrint-Gewöhnung unter verschiedenen Graden visueller Hilfe. Sechzig Collegemädchen dienten als Versuchspersonen. Sie wurden in vier Gruppen zu je fünfzehn eingeteilt: eine Kontrollgruppe (0) und drei Experimentgruppen (1, 2, 3), alle mit gleichen Fähigkeiten, bestimmt durch den Alphatest. Allen Versuchspersonen wurden während des Nachspürens die Augen verbunden. Die der Gruppe 0 hatten keine visuelle Hilfe, währenddem die der Gruppen 1, 2 und 3 das Labyrint 5, 10 bezw. 15 Sekunden vor jedem Versuch sehen konnten. Als Kriterien der Lernfähigkeit nahm man die Anzahl der Versuche, der Fehler, und der Sekunden an, die notwendig waren, um die Fähigkeit zu erreichen, in drei Versuchen zweimal fehlerfreie zum Ziel zu gelangen. Der Grad der Zuverlässigkeit hinsichtlich der Raten wurde in den Experimentgruppen stark erhöht (0.75, 0.87 und 0.87 für 5, 10 und 15 Sekunden Exposition; 0.48 für

die Kontrollgruppe). Die Methode bestand darin, die Fehler in geraden und ungeraden Versuchen zu vergleichen. Visuelle Exposition ergab ferner eine beträchtliche Reduktion der Lernschwierigkeiten. Wenn der Schwierigkeitsfaktor ausgeglichen worden wäre, würde der Zuwachs der Zuverlässigkeit visueller Hilfe zweifellos noch grösser gewesen sein. Der Author glaubt, dass ein auf das Optimum der visuellen Exposition geeichtes Hochrelief-T-Labyrint ein befriedigendes Experiment für die Labyrintgewöhnung der Menschen ergeben würde, dessen Ergebnisse ziemlich gut mit Massen der Intelligenz korrelieren durften. Für allgemeine Bedürfnisse wäre es wünschenswert, mehrere verschiedenartige Tests zu haben, die untereinander geringe Korrelation hätten, von denen aber jeder eine hohe Korrelation mit einem der Intelligenzamasse hätte.

Peterson und Allison

EXPERIMENTAL METHOD IN PSYCHOLOGY*[1]

F. C. Bartlett

The position which Psychology occupies today in relation to the other biological sciences, both theoretical and applied, is due directly to its adoption of experimental method. Yet if the psychologist is asked to point out any single unshakeable discovery of first-rate psychological importance, based directly and wholly upon experiment, his attempts to answer the question are always regarded as unsatisfactory. Moreover, it is still common to find the most skilled and notable psychologists beginning their careers with an enthusiastic and active belief in experimental method and, as they develop, receding farther and farther from the laboratory and becoming more and more attracted by the claims of systematic general theory. Two reasons are assigned by everybody for these facts. Each is valid in itself; neither is helpful nor reassuring. They are, first, that experimental psychology is a recent growth, and, secondly, that all psychological problems are discouragingly complex. It would be a bold thing to maintain that any method used in scientific investigation is fruitful in direct proportion to its age, while the fact that all psychological problems are complex is no justification for attacking them by theoretical methods which are of all ways the most likely to lead to over-simplification. It therefore seems worth while to attempt a survey of the use to which experiment has been put in psychology, in order to show what experiment has done, is doing, and what there is reasonable ground for believing that it yet may do for the advancement of psychological study. Such a survey may help to explain the curious fact that the very method which, more than anything else, has raised psychology to something like equal status with related sciences, is still regarded with suspicion both within and without the borders of the subject.

As everybody knows the earliest experimentalists in psychology were physicists and physiologists, most of them with a strong bent towards philosophy in their outlook. They set up a standard which,

---

*Received in the Editorial Office, November 4, 1929.

[1] Address by the President of the Section to Section J.—Psychology, British Association for the Advancement of Science, South Africa, 1929.

in various ways, has cramped and confined experimental psychology ever since. When a physicist approaches a problem in which he has to state how a stimulus effects any kind of response, he is bound to lay the burden of explanation upon the stimulus. If this can be simplified, controlled, varied determinatively, resulting differences of response can be observed and explained. The stimulus also tends to be treated as outside the response itself. When the physiologist approaches the same type of problem, his emphasis is equally bound to be mainly upon the response side of the relationship. This he perfectly legitimately isolates as far as possible, trying to state its characteristics in terms of the functions of the mechanism immediately concerned.

These are precisely the type of problem that the psychologist has most often attempted to carry over into his own field. Let us consider one point which grew out of the early experimental work on psychophysical methods. Fechner believed that all the content of experience, from the relatively simple sensation to the most complex reasoning content, was definitely measurable. Mainly as a result of his own observations, but aided also by the earlier work of Weber, he thought he could demonstrate this experimentally by showing that stimulus intensities and sensation intensities are related by a definite principle, the value of the unit sensation varying with the modality of stimulation. In order to establish the zero point for a sensation of given mode, and also the series of just noticeable differences piled up from this point by increase of the stimulus, he carried out a wonderful and patient series of experiments. He employed and somewhat improved Weber's method of "limits"; he developed the method of "mean error," but he relied mainly upon the method, already proposed by Vierordt, of "right and wrong" cases. Fundamentally this method consisted in the presentation to an observer of a series of stimulus variables, each variable being presented a number of times over together with some standard magnitude, the order of presentation of the variables being haphazard. Thus for each stimulus variable a number of judgments were obtained upon its equality with, or difference from, the standard. As might have been expected, the observer was often doubtful and said so. Now, argued Fechner, if these doubtful judgments were resolved by an ideal observer, they would, in the long run, be equally divided on both sides of the alternatives which confront him every time he judges under these experimental conditions. Consequently in determining the stimulus value which gave a sensation difference threshold he adopted the plan of dividing the doubtful cases equally between "right" and "wrong" cases.

This procedure was speedily and acutely criticized by G. E. Müller. Müller pointed out, among other things, that the judgment in these cases is not simply a function of the stimulus on the one hand and of the sensitivity of the observer on the other hand. It is determined also by the *precision* of the observer, and this will affect the actual significance of the doubtful cases. But he did not go all the way with this notion; he thought that statistical manipulation could give him a measure, both of precision and of sensitivity.

Thus grew up a controversy which has proceeded voluminously ever since, though at the moment it seems nearly to have worn itself out. Broadly, four plans for dealing with these "doubtful" judgments have been proposed and followed:

(*a*) to divide them equally between "right" and "wrong" cases;

(*b*) to ignore them;

(*c*) definitely to instruct the observer to "guess" when he is uncertain and then, if doubt still persists, to lump all cases together with "guesses of greater" and "guesses of less" as constituting an area within which stimulus variations have no corresponding sensation differences;

(*d*) to prohibit the observer from being doubtful.

Now all these devices ignore the very points which are psychologically the most interesting and important. First, they tend to treat each judgment in the series as equally and independently significant, the function of the immediate stimulus. This is certainly wrong. For example, Wundt demonstrated long ago that a given judgment may express, not the immediate effect of its stimulus, but the cumulative result of a series of preceding stimuli and situations, no one of which at once issues in a characteristic overt response. Again and again the status of a pronouncement in a series has been shown to depend on its order of presentation within that series. Secondly, no judgment of this type is the expression of a simple stimulus-response situation, but of a stimulus-attitude-response situation. To demand guesses and to prohibit doubt both alike determine an *attitude* of observation which spreads over the whole experimental situation, affecting judgments which are assigned certainty just as much as the others. During a few years preceding the war a number of extremely interesting experiments were carried out, particularly by Dr.

Fernberger, upon the effect of "attitude" in determining judgments obtained by psychophysical methods. For some reason these have never been adequately followed up, but they represent what I take to be a genuine problem of the experimentalist in psychology. In this case the question is: How are the judgments obtained in the various psychophysical methods determined? It is not enough to correlate immediate stimulus variations with immediate characters of response, or simply to describe the immediate response mechanism that is brought into play. When an observer enters into an experimental situation he brings with him propensities, tendencies, preformed organized systematic modes of response, the performed cumulative organized effect of a mass of past discriminations. The stimulus, the situation that is presented, hits off some of these. They appear in him as an "attitude," and it is under the active control of this that he makes his responses. Only when we know more about how this is set up and about its precise effect upon the responses made, can we safely give to the latter the necessary weighting which makes their statistical treatment genuinely significant. I think that the psychophysical methods, studied from this point of view, will yet yield some enormously important results.

If the physicist in his approach to the stimulus-reaction type of problem tends to treat the stimulus regarded objectively as the main point of interest, the physiological method of approach is equally bound to concentrate its attack upon the immediate functional mechanism. It is, of course, foolish to argue, as is often done, that the physiologist is absolutely confined to a study of local response mechanisms, and that it is a matter of indifference to him whether they are in or out of their wider organic setting. But he does rightly lay the chief emphasis upon these and for obvious and just reasons shrinks from speculation about central processes until he knows as much as can be learned about the peripheral functions and their mode of operation. From the earliest days up to the present a great part, perhaps the greatest part, of experimental psychology has been concerned with special-sense reactions. Yet the psychologist has never staked out clearly any mode of investigation or characteristic problems which are especially his own in this field. In this he has again been over-influenced by the great founders of his science. Consider the work of Helmholtz and of Hering on visual reactions. The former, with his very predominantly physical outlook, laid the burden of explanation always upon what the stimulus and variations of the stimulus can do. Only when he got into tremendous difficulties, as with simultaneous contrast, did he invoke processes, not directly visual, seated in the observer himself; and then, at once, his explanations became confused and unconvincing. Hering, being far more dominantly physiological, turned to a study of the immediately functioning mechanism. He, however, seemed to set the stage for much so-called psychological interpretation in later days, by elaborating a sort of speculative physiology in his theory of anabolic process within a local sense organ. When a psychologist takes up a problem of special-sense reactions, his whole training and outlook force him to lay due weight upon the wide variability of such response as it takes place in its normal organic setting. If he has his eye upon the physiological mechanisms this variability constantly forces him back beyond the immediate peripheral functions to more central processes. Then, as again and again in the history of the subject, he is tempted to draw fanciful diagrams of the brain, laid out neatly into all sorts of areas of function; or, as is perhaps more common nowadays, he relapses into wholly speculative physics and physiology of the central nervous system. This is only trying wildly to do what the real physicist and physiologist are most chary of attempting, except as a kind of game, and it is no wonder that many of the more scientific psychologists are somewhat discouraged, while other people find it difficult to take the psychologist seriously.

Yet it is not the psychologist's view of the problem, but his handling of it, that has been faulty. Suppose we are investigating some specific sensory threshold. The response of our observer will be determined by a number of groups of factors. There are the physical characters of the stimulus: its intensity defined physically, its duration, often its medium of conduction to the sense organ concerned, relevant facts of its physical and chemical structure. There is the absolute sensitivity of the local responding physiological system, though this can never be absolutely measured short of cutting it out of its organic setting, which is just what the psychologist must not do, though the physiologist certainly may. There is the order of presentation of the stimulus in its series, which raises problems already touched upon in connection with the psychophysical methods; and the correlated physiological questions of the state of adaptation at the moment of the sensory system. Operating over and above and through all of these are the tendencies, attitudes, moods, intellectual and emotional habits of the observer, the states variously characterized as states of confidence, hesitation, doubt, timidity, assertiveness, certainty. An image, flashing out suddenly at a given point, may change the whole character of a response and of succeeding responses. So may the verbalization, or formulation, of a judgment. These so-called higher mental processes are precisely the psychologist's main concern. They can be experimentally set up and controlled to a large extent, while the other factors are kept relatively constant. It is our business to show how they are set up, and how they then powerfully determine reactions within the special sense field, even in the simplest cases of such reactions. It is of no use simply to say that there *are* these determinants, and then to take refuge in a fruitless hypothetical physiology of the central nervous system. Perhaps the time may come when the bulk of sense psychology will be swallowed up in physiology, but that time has not come yet, and if we act as if it has we gain little but deserved suspicion from other scientists.

Already there are a few experimental studies of special sense problems from this point of view. They are less frequent and less thorough than they should be, and this is because the psychologist has generally been content to follow his more physically and physiologically minded predecessors, instead of envisaging the problems of sensory reactions for himself. When the psychologist studies a special sense response it is his business to try to show how that response, carried out in its normal organic setting, is being determined directly by facts other than those of the immediate sensory mechanisms. Cues other than those of the special sensory stimulus are operating through response systems of a higher, or more complex, order than those of the directly excited sense. At present psychological names have to be used for these response systems, and we have to show how they come into play and what they do.

It should, therefore, be particularly interesting to turn to the experimental attack upon the higher mental processes. I propose to take as typical the study of the responses called "remembering" and "recognizing."

The experimental investigation of memory is dominated by the work of one man who is commonly supposed to have been a great benefactor of experimental psychology, but who, in spite of his impressive work, seems to me to be the errant leader of a very sheep-like flock. In 1885 Hermann Ebbinghaus published his new program for experiments on memory processes. Already since 1879, he had been studying his own modes of recall by methods which were at that time novel. The publication of his results settled the direction of flow of the main stream of experiments on memory from that day to this. As everybody now knows, his great innovation was the use of nonsense syllables for memorizing. He claimed for these four great advantages over any other type of material: they are simple; they are homogeneous; they can be indefinitely combined, but in all combinations the material remains essentially on the same level; they "admit of quantitative variation which is adequate and certain."

It is fairly easy to show that not one of these four claims is in fact sound, but for the present a single fundamental consideration is enough. It is always urged that nonsense syllables are the best material to use because they are "simple." That appears to mean something like this:

Suppose we are investigating normal taste and smell reactions. We know that generally these two are inextricably combined, so that much that we call taste is really smell. However, if we are interested in finding out how the end organs of taste react, we can experimentally cut off all olfactory reactions and see what happens. We then learn something true and important about how the end organs of taste behave when they are subjected to certain specific conditions. It is more than hazardous to assume forthwith that they behave in precisely that manner when the olfactory sense is alive and working. But there is, no doubt, a sense in which we can say that the taste reactions thus experimentally studied are "simpler" than those of normal everyday life. Now take the analogous case of remembering. A part of the total process is obviously the initial learning. If we use ordinary combinations of meaningful words or forms there may be a mass of associations, in some respects peculiar to each subject who submits to experiment, set up during this initial period of observation. By the use of nonsense syllables we cut off all, or most, of these thronging associations, just as in the other cases we cut off the olfactory sensations, and so we find out what pure and simple remembering can do.

The argument is feeble. We do not isolate the taste experiences by simplifying the stimulus, but by cutting off, through extirpation or temporary paralysis, those other reactions with which they are normally integrated. We can operate on the olfactory nerves in this way if we desire, but there are no specific memory nerve endings or nerve centers, or if there are we do not know them. Moreover, even if we could do this, the experimental psychologist would be untrue to his pretentions if he were satisfied with this alone. He pretends to deal with the intact organism. Suppose we do simplify enormously our stimuli, or our experimental situations and then confront our intact subject with these. All that happens is that he is faced by such an odd and unusual state of affairs that he is forced to mobilize all his resources and make up some novel reaction *ad hoc.* This is, of course, exactly what happened when the nonsense material was used. The early experimenters, being in other ways exceedingly good psychologists, saw this at once. Their way out was curious. They proposed that every subject must have long-continued practice with nonsense syllables before any of his results should be allowed to count. That is to say, having taken immense precautions to simplify experimental material and methods, we must then take equally great precautions to simplify, by rendering familiar, the highly artificial and complex response which we have thereby set up. The way in which this sort of method has run rampant over the whole of the laboratory psychology of the higher mental processes is distressing to anybody who wishes to adopt an experimental approach to this field of research. It is, perhaps, hardly an exaggeration to say that nine-tenths of conventional laboratory psychology —outside of experiments on the special senses—consists largely in showing how habits of response may be set up towards odd and out-of-the-way situations, and makes little contribution of moment towards the solution of any other problem.

This is the old difficulty in a new form, the exaggerated respect for the stimulus or the situation. The psychologist is studying the complex responses of a highly developed organism and how they are determined. They have been called forth to meet the claims of a very unstable and varying objective environment. They are, no doubt, in many ways more stable than that environment. But if the environment is violently simplified it is mere superstition to trust that they also get simplified in a corresponding manner. They become different, but are just as likely to become yet more highly complex. Stability of determination, not simplicity of structure in objective determining factors, is what we need to make our experiments convincing. Stability of determination is compatible with complexity, and even with considerable variation of objective determinants. Somehow an experimental method has to be developed which recognizes this fact.

I turn for a moment to experimental work on "recognition" merely to bring out one further point of method. Since very early days, an enormous amount of work has been done on this topic. It has issued in five or six different theories no one of which can claim finality. The diversity of this result is due to different causes, but one is perhaps particularly important. There is a strong tendency, when any complex response like recognition is being studied, to attempt to draw a ring round it and to seek its explanation within these imposed limits. Thus the explanation of recognition is sought in something that happens at the moment of recognition. This is surely wrong. An object or event may be recognized or not largely on a basis of how it was reacted to in the prior perception. A sound, for example, may be heard: it will not be recognized unless it is so listened to that it possesses qualities, characteristics, a setting, and a significance. To the persistent study of a complex mental response as if its psychological explanation must be found inside an imaginary circle that encloses it, much of the disrepute into which experimental psychology tends to fall may be traced. It is perhaps the last and subtlest form of the outworn "faculty" psychology.

I said at the beginning that the early experimentalists in psychology were physicists and physiologists with a strong bent towards philosophy. If, in the history of the subject, experimental psychologists have shown themselves too submissive to physical and physiological methods, it is even more true that they have often pursued philosophical ideals. This pursuit is in full cry still. A very brief consideration of current movements which originate in the laboratory will illustrate this point.

There is probably no contemporary movement in psychology which has more profoundly influenced psychological thought in English-speaking countries than the so-called *Gestaltpsychologie.* I have, as every experimental psychologist must have, a very great admiration for the brilliant work of Wertheimer, Köhler, and Koffka. It has shed much new light on old problems, as well as a good deal of old light on new problems. It starts specifically from experiments upon visual perception, and its primary method is that of phenomenological description. When we are presented with a perceptual situation what is it that we experience? The answer is inevitable and is one which, from this point of view, has, I think, always been given. We cannot describe our experience in this sort of situation as a mosaic of tiny bits each corresponding with its isolable part of the stimulus or situation. The blue sky which is seen is not, as Köhler says, made up of an infinite number of blue sensation units, but is seen as a continuous blue expanse. The moving dots and lines, in Wertheimer's experiments, are seen, not as stationary points in temporal relations, but as a unitary and indivisible movement. Sometimes the experiments are more behavioristic, but still it is the attitude of phenomenalistic description that determines their interpretation. The animal that has been trained to react positively to *a* and negatively to *b* at once reacts positively to *b* if *a* is removed and *c* is introduced bearing that relation to *b* which *b* had to *a*. This must be because the initial reaction was not to *a*, or to *b*, and not to *a* and *b* and a relation between *a* and *b*, but to a total indivisible situation only to be described as *a-b-in-relation*. Again, an animal is set a complex problem and typically achieves the solution suddenly; for what we call the "correct" solution is a reaction to the total situation as built, or figured, or formed.

Thus a fundamental psychological question tends to be: "What is the nature and what are the characteristics of these indefeasible forms or patterns which stand over against all our reactions, compelling them to be as they are?" In our answer we can easily slip into the persistent error of over-emphasis of the objective side of the situation-response problem. I do not say they do this, but in his doctrines of physical *Gestalten* Köhler comes very near it. Or again, looking to the response side, we may try to build up inside the responding mechanism a complex system, somehow corresponding to the integrated phenomenal situation. Then, as in Köhler's theories of ionic concentration within the central nervous system, we are almost sure to slip into sheer speculative physiology. Finally the phenomenological attitude seems bound to issue in a comprehensive theory about the nature of the world as we know it, rather than in a scientific study of the determination of human response. It is the latter alone which is truly amenable to experimental treatment.

There is another contemporary movement, also originating in Germany, much less widely influential in other countries at the moment, but likely to attract more and more attention. This springs from the work of Professor E. R. Jaensch. It also has a definite experimental basis. Jaensch experimentally discovered a type of imagery which seemed to lie somewhere between the after-sensation on the one hand, and the genuine memory image on the other. I cannot attempt to describe his extremely keen investigaton of this *eidetic* imagery, as he called it. At the moment the main point is that he considered it demonstrable that a proneness to eidetic imagery is correlated with a number of other reaction tendencies. He elaborated a theory of the two-fold division of all human subjects into integrate and disintegrate types. The integrate is the artistic, synthetic type, taking everything as a whole and having an inevitable accompaniment of persistent marked temperamental qualities and tendencies. The disintegrate is the scientific, analytic type, tending to split up presented situations and to deal with them piecemeal, and he also has his invariable, persistent, accompanying temperamental character. Of course the integrate and the disintegrate in their most marked forms are the extremes of a very wide range and pass, the one into the other, by very small shades of difference; but there remain these two predominant types, by a study and understanding of which all the problems of human reaction, in every field whatsoever, are to be finally explained.

Now this view does certainly seem to be biological in bent, and it is being explored throughout by experimental study. Moreover, it strives, and, I think, successfully, to avoid that artificiality which, as we have seen, hangs over conventional laboratory methods for the investigation of the higher modes of human response. It rightly treats our problems as problems of reaction tendencies, of their determination and their grouping. But it does seem to be rather in a hurry with its sweeping generalizations. Here are some extremely interesting observations on imagery. Why should we hasten to fashion from them a key to all the most difficult problems of our science? It is what the psychologist seems everywhere prone to do. A departmental investigation is made to carry the weight of a comprehensive system. This, it seems to me, is the philosophical bias at work.

In contemporary English psychology there is only one complete and world-explaining system of this sort, and that is the system which has been developed by the keen and penetrating work of Professor C. Spearman. Here again we are invited to begin with experiment, not merely upon perceptual processes, not merely upon imagery, not merely upon more complex intellectual responses still, but upon any psychological problem whatever. Wherever we begin we shall speedily find illustrated the working of the few immutable laws upon which all mental structure, and it may be the very Universe itself, are built. They stand indeed at the portals of our science.

Know them and everything is plain; ignore them and all is confused. They are conceived, not as tendencies serving the ends of biological adaptation, but after the fashion of physical principles of universal scope describing the inevitable structure and frame of mental life.

That Professor Spearman has done more for the development of several fields of psychological research than any other living Englishman cannot be disputed. His contributions to the development of psychological statistics, and his work in the field of the investigation of intelligence must give him a permanent place in the history of the subject. Yet the notion of experiment mainly as a tool for laying bare and illustrating the operation of a few basic and fixed laws must seem to the biologically minded investigator very unsatisfying. Having found a scheme which seems to set the results of certain experimental investigations in order, why should we, with all the wealth and variability of human response as our subject, try to fit the same scheme in everywhere? Surely to do this either our scheme must be so broad that its explanatory value in relation to particular concrete problems is bound to be very remote, or else it must be ruthless in its dealings with fact. In writing some years ago of his "theory of two factors," Professor Spearman said: "It would seem as if psychologists have now got definitely to accept the Theory of Two Factors; it becomes a Bed of Procrustus into which all our doctrines must somehow be made to fit, even though the so doing may at times involve a not unpainful surgical operation upon them." I admire the courage that prompts such a statement, but I find it difficult to square this attitude with that scrupulous regard for the immediate facts which must mark the experimental biologist.

No survey, however scrappy, of contemporary movements in experimental psychology can be satisfactory without some reference to Behaviorism. Of all the movements, this is the one which is most thoroughly experimental, alike in its methods and in its formulation of problems. It has laid firm hold on the point of view that experimental psychology is an investigation of the conditions determining high level biological reactions in animal and man. It is so round in its denunciation of philosophy that its excessive readiness to systematize its own principles of explanation is amusing. We can see this readiness in the haste with which it has exalted the principle of "conditioned reflex" into an all-embracing explanation, though many of the problems of human response concern the emergence of new effector functions and conditioned reflex has nothing to do with this, and though conditioning at the human level is excessively speedier and often far more stable than anything that has ever been experimentally observed. We can see it, also, in the Behaviorist's dogmatic assertion that the development of consciousness within any type of biological response never makes any difference in subsequent response. Such dogmatism is only another instance of the experimental psychologist's fatal proneness to run beyond his data. It is explicable in the light of a study of the origins of Behaviorism, for it was by the adoption of Behavioristic methods alone that the investigation of animal response below the human passed from the anecdotal and analogical stage and became genuinely a part of biological science. But to push the principles involved into the whole of human psychology is just as bad as to carry out some departmental investigation into perceiving, or imaging, or thinking, or some sensorial function, and then to use the results forthwith as a master-key to all the problems of human determination.

Some of the reasons why experimental psychology has often attracted unfavorable criticism and failed to hold its students should now be clear. In work on the special senses it has frequently attempted to deal with problems that the physicist or the physiologist with their specialized training could solve more satisfactorily. In dealing with the higher mental processes it has been over-impressed with the necessity of standardizing objective situations and has constantly proceeded as if the simplification of a stimulus were equivalent to the isolation of a response. Persistently it has shown unnecessary readiness to build upon specialized investigations wide systems which pretend a finality and universality that they do not possess. I believe the time has now come for pushing these criticisms vigorously and for attempting to meet them in practice. It could hardly have come much earlier. After all, whatever the limitations of his outlook, it may fairly be claimed that the experimental psychologist has done more than anybody else to keep alive the interest in special sense problems. As for the work upon relatively more complex processes, a new and struggling science was almost bound to imitate methods already fully established in other fields and to exploit them as far as they would go.

By its shortcomings and failures, as well as through its successes, experimental psychology has been attaining the rank and outlook of a biological science. I will conclude this survey by saying as definitely as I can what I take this to mean.

Experimentalists everywhere are directly concerned with a study of the conditions under which the observable results that interest them can be shown to occur. In psychology the experimentalist is attempting to find out the conditions of the various reactions or modes of conduct that make up the lives of animals and human beings. At the human level most of these reactions appear to be accompanied, and in part determined, by some form of *content*: by sensations, images, judgments, trains of reasoning. Sometimes the psychological experimenter prefers to put his problems directly in terms of these, of how they occur and how, when they have occurred, they may, as conditions, influence other forms or instances of content. Whether he deals with these questions or not there is one thing that is always held to distinguish his point of view from that of related scientists. He is primarily interested in the intact organism or the intact mental life. For the most part he cannot cut out partial responses, or special forms of content, and consider how they are conditioned in the absence of all the rest. If he could do this it would be foreign to his essential problem to make the attempt. He may, indeed, gain help from a study of instances in which certain types of response have been lost, leaving still a very complex organism, or mental life, to effect a readjustment of what is left. This, however, is clearly a very different matter from cutting out the partial response and studying that by itself.

The tendency to try to treat partial or specific forms of response each as much as possible by itself has, as I have indicated, persisted from pre-experimental days, and has done much to render experiments in psychology often trivial, uninteresting and highly artificial. The organism in which the psychologist is interested does not normally react with one sensory modality only, or with emotion untinged by information, or with impulsive activity free from interference by non-impulsive factors. All this might be unimportant, if by experiment the psychologist could force his subjects to adopt such simple modes of reaction. The psychologist would indeed be barred from making that direct passage from the laboratory to the world at large which he is very fond of making, but that would not be an unmixed evil. The situation is, however, more difficult than this. The subject of a psychological experiment always retains the possibility of making more reactions than the one that is being specifically studied. No matter how honest and painstaking both experimenter and subject may be, these "other" reactions will play their parts in any results obtained. To try to shut up a subject to a purely visual or a purely auditory reaction, to pure perceiving, pure remembering, pure imagining and the like is, in fact, only to force him to utilize indirect cues so different from those which he constantly employs under normal conditions as, in many cases, to make his total reaction a genuinely novel one. The reaction has become artificial, that is, not merely because the range of its conditions is restricted, but because, with as wide a range of conditions imperfectly under control as ever, many of these are specific to this special experimental situation. The perfecting of objective control does not reduce the subject to a single sense, or a single cognitive or emotional process, but, leaving him as complex a reactive agent as ever, forces him to make up on the spot a type of total reaction fitted to this special environment.

Holding all this firmly in mind, what conclusions can we draw? First, it follows that the experimental psychologist must claim that for the present, and perhaps for always, he is as much clinician as experimenter. He has not merely to arrange conditions and record results. There seems to be a notion abroad that there is so much uncharted ground in psychology that an investigator can do anything he pleases, and so long as he observes everything possible, his results are bound to be significant. This is utterly false. His observation is definitely that of a man with a problem, and generally also with a personality in view; and it is by consequence almost glaringly selective. He is not alone among experimenters in this respect. From a reading of the theory of the matter one might be tempted to suppose that the best experimenter would, once the experiment is arranged, be merely a rather complicated and delicate recording and calculating device. Those who have a reputation for brilliant experimental work in any field singularly fail to impress this character upon the intelligent and sympathetic onlooker. Anybody who, by experiment, is going to discover anything important about the determination of human reactions must first have developed a certain character of human reaction for himself. If this is to be used against him when he claims validity for his discoveries, it is a sort of stone he can return with some effect, whoever his opponents may be.

Yet it is extraordinarily important that the experimental psychologist should not be exclusively concentrated upon the particular reaction which he is specifically studying. Just because it is the intact subject, the intact organism, that we are concerned with, the conditions of any reaction are apt to branch widely. The problem for us, for example, is not to find out how the eye sees or the ear hears, but how the animal and man do. No doubt we can answer this problem only in an imperfect way, but it takes us no nearer perfection to cut the ear or the eye out of the man. This is true with increasing force as we go higher in the level of response. Indirect cues are neither to be ignored, nor to be cut out, but definitely to be studied.

Secondly, no experimental psychologist must profess, with unvarying belief, the dogma of constancy of objective conditions. If, biologically speaking, human reactions had been built up to meet a series of unchanging environments, emphatic insistence upon rigidity of conditions would be justifiable. Obviously, they are not so built. So far as the psychologist is concerned, many of the most important characters that dominantly set the course of our reactions belong directly to the organism with which he is dealing, to its immediate and remote past history and to its present specific and general state of adaptation. These intra-organic, or intra-subjective factors may be more diversified by rigidity than by variation of outward circumstance. Of course, I do not argue that the experimental psychologist need not be seriously concerned with constancy of experimental procedure. He must arrange this as carefully as possible, taking advantage of whatever may be known as to the order of importance of the elements of any complex situation with which he is concerned. But he should never hesitate to break this constancy of procedure if his psychological judgment or insight assures him that a diversity of outward circumstances will best secure a relative stability of attitude in his observer. To the experimenter who is not a psychologist this claim may appear arbitrary and arrogant. There is, so far as I can see, no help for that.

In the third place, the position which I have stated carries with it that the experimental psychologist, at the end of his studies, has to be satisfied with indicating trends, directions, proclivities rather than dogmatic laws. His phenomena are essentially biological, in process of development, displaying no hard and fast boundaries anywhere. He may formulate dogmatic laws, and use experiments as imperfect illustrations, but this is the wrong order of things, though it has been by far the commonest and is still the easiest.

Finally there is the question of the relation of the results of specific experimentation to the claims of general systematic theorizing. It should be clear that I am not for one moment for the haphazard experiment that has no idea, no broadly formulated problem, behind it. Also I would condemn as heartily as anybody that scatter of descriptive results, uncoordinated, unsystematized, which is common in many directions nowadays. We must explain our results and not merely collect and exhibit them. Yet I would urge that when we have, for example, satisfactorily stated the conditions of some particular perceptual reaction, we have no more right to pronounce magisterially upon a complex problem of reasoning than a physiologist who has studied respiratory functions has to pretend at once to clear up the secrets of spinal reflexes. No doubt the physiologist would never for a moment attempt to do this, but unfortunately it is not so easy to answer for the pretensions of the experimental psychologist in a like case. It may even be that all our specific studies will lay bare common broad principles of the determination of response. Even so, the broad principles are not the explanation of the specific problem, and for whatever they may be worth, before we erect them into a comprehensive system we must have the specific problems widely and patiently worked out.

With this outlook of mingled boldness and caution I believe that experimental psychology will prove adequate to its task of building up a sound scientific study of complex response in animal and man.

University of Cambridge
Cambridge, England

LA MÉTHODE EXPÉRIMENTALE EN PSYCHOLOGIE

(Résumé)

La psychologie expérimentale est le résultat de l'oeuvre des physiciens et des physiologistes dont la plupart ont possédé une forte inclination philosophique. Ceci a entravé son développement, en premier lieu en faisant que les psychologistes appuient trop sur les explications basées simplement sur une étude des variations du stimulus, en deuxième lieu en causant beaucoup de speculations fantastiques sur la physiologie du système nerveux central, et en troisième lieu en donnant lieu directement aux systèmes rigides d'explication psychologique. On illustre ces points en citant des études de méthodes psycho-physiques, de psychologie spéciale des sens, d'investigation des processus de memoire, et une consideration de quelques systèmes psychologiques contemporains.

Un psychologiste est un biologiste qui prend comme champ de ses études les conditions des réactions de l'organisme intacte ou de la vie mentale intacte. C'est un clinicien aussi bien qu'un expérimentateur. Il doit éviter l'isolement excessif des réponses qu'il étudie; il doit éviter le dogme de la constance des conditions objectives; et puisque ses problèmes sont biologiques son but n'est pas d'établir des moyens définitifs d'explication, mais d'arriver à des principes explicateurs qui indiquent des tendances, des directions, des inclinations plutôt que des lois dogmatiques.

BARTLETT

EXPERIMENTELLE METHODEN IN PSYCHOLOGIE

(Referat)

Experimentalpsychologie ist aus der Arbeit der Physiker und Physiologen herausgewachsen, von denen die meisten eine stark philosophische Neigung hatten. Diese Tatsache hat ihre Entwicklung gehemmt, erstens dadurch, dass sie die Psychologen verführte zu grossen Nachdruck auf das Studium der Reizvariationen zu legen, zweitens dadurch, dass sie zu vielen kapriziösen Spekulationen über die Physiologie des zentralen Nervensystems führte, und endlich dadurch, dass sie zu harten und festen (fast) Systemen psychologischer Erklärung führte. Diese Punkte werden unter Bezugnahme auf psychologische Studien, auf specielle Sinnespsychologie, auf die Untersuchung von Gedächtnisprozessen und unter Betrachtung einiger zeitgenössischer Systeme illustriert.

Ein Psychologe ist ein Biolog, der als Gebiet seiner Studien die Reaktionsbedingungen von intakten Organismen oder des intakten Geisteslebens gewählt hat. Er ist sowohl ein Kliniker als ein Experimentator. Er muss unangemessene Isolation der Reaktionen vermeiden, die er untersucht; er muss sich vom Dogma von konstanten, objektiven Bedingungen frei halten; und da seine Probleme biologische Probleme sind, muss sein Ziel sein, nicht endgültige Erklärungsschemata festzulegen, sondern zu erklärenden Prinzipien zu kommen, die eher Tendenzen und Richtungen, als dogmatische Gesetze anzeigen.

Bartlett

"SPONTANEOUS ACTIVITY" OF MICE*

From the Laboratory of General Physiology, Harvard University

T. J. B. Stier†

Movements of animals are sometimes thought to be initiated solely by stimuli impinging upon peripheral sense receptors. If, however, stimuli of equal intensity are applied at regular intervals, it is observed that, in general, responses of equal intensity are not obtained. The kind and the extent of this variation in response is not the same in all cases. Furthermore, if the experimenter endeavors to exclude all stimuli from the environment, he finds that bodily movements still occur in the animal under observation; movements of the whole organism occurring under such experimentally controlled conditions are called spontaneous movements. We shall deal with experiments designed to reveal the character of a mechanism which may be responsible for "spontaneous movements," and thus for one type of "variation" in behavior. Speculations as to the nature of a possible scheme are presented, not as embodying a final or definite theory, but as essentially a plan for future experimentation which grows out of the observations.

Movements in Newborn Mice, as Related to Temperature

I

For these experiments I have employed young mice known to be genetically homozygous for the albino character, having been inbred brother to sister for 30 generations, originating from a single pair of the twenty-fifth generation of brother-sister matings. It has been found [MacDowell and Lord (18), Crozier and Pincus (10)] that such genetic homogeneity is absolutely essential in order to obtain experimentally reproducible and significant results. Animals two days old were employed because it was thought that the conduct of these young individuals would be free from the complications of conditioned reflexes, undoubtedly occurring to a greater extent in older mice. It was found [Stier and Pincus (26)] that these young mice were practically poikilothermic, and could therefore be employed for tests intended to relate features of activity to temperature. These experiments would then serve as a kind of test for deciding whether the organism has control over its own activity—free will—or whether it is simply a physico-chemical machine which must increase its activity according to well-known laws of chemical dynamics as the temperature is increased.

---

*Received for publication by W. J. Crozier of the Editorial Board, November 20, 1929.

†National Research Council Fellow.

At first it was thought that the experimental technique should be very "simple." The mouse should be allowed to move about, unattached to any apparatus which might add to the already complicated inner mechanism. I would, so far as possible, control all external variables—temperature, humidity, noise, light—and then record the activity on a kymograph by actuating a signal magnet whenever the animal moved. An apparatus containing 15 signal magnets has been devised by Pearl (21) for registering all the kinds of movements made by *Drosophila*.

When a two-day-old mouse was placed on a piece of cork in the bottom of a beaker, the temperature of which was controlled by immersion of the beaker in a water bath, records of movement were obtained, as in Figure 1.

If we let the mouse move around on the platform of such an apparatus as is illustrated in Figure 5, we obtain an equally unintelligible record. A sample is reproduced in Figure 2.

So far, spontaneous movements seemed to occur without sufficient regularity for making precise measurements at different temperatures.

It was noticed, however, that two-day-old mice left free to move about on the recording platform do not always come to rest in the upright position after a period of spontaneous activity. These muscularly feeble animals sometimes go into a state of quiescence, lying ventrum down, sometimes on their sides, or even on their backs with their feet in the air. It seemed obvious, therefore, that afferent stimuli from proprioceptors in the muscles and tendons of the legs would be different in kind and amount in each of these positions. Moreover, impulses involved in reflexes such as those dealing with righting of the animal would activate it much sooner when in any of these positions than if the animal were constantly held in an upright position. It is due to the explosive all-or-none character of the initiation of the periods of activity that the toppling-over movements occur.

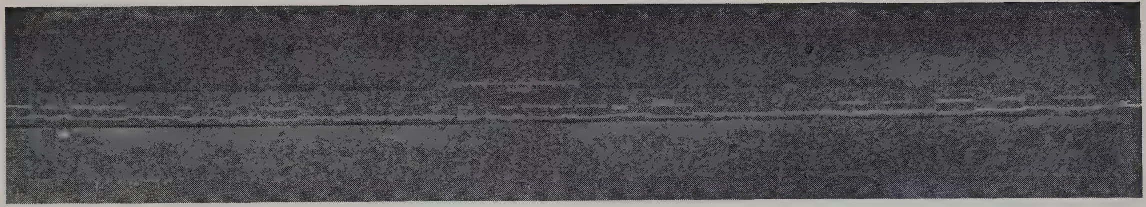


FIGURE 1

A Record of Movement and Quiescence of a Two-Day-Old Mouse

The animal was free to move about on a piece of cork in the bottom of a beaker. Two signal magnets actuated by pressing appropriate tapping keys marked the beginning and end of activity. There is no regularity to the cycle of activity.

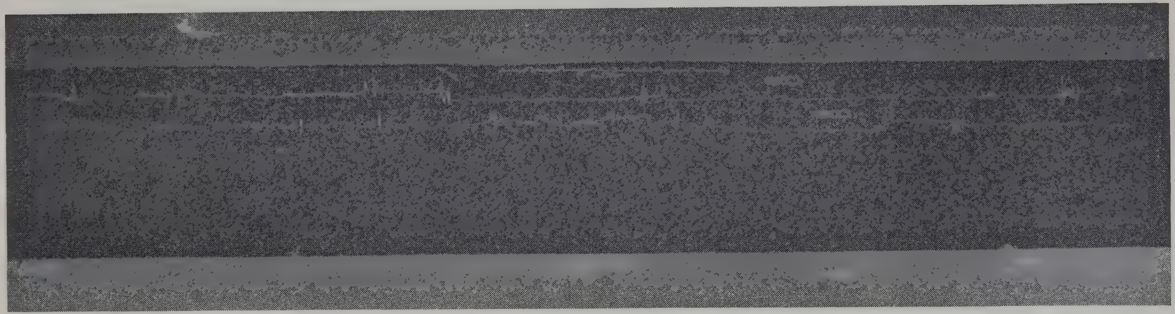


FIGURE 2

Two Records of the Movements of a Two-Day-Old Mouse Obtained at 21.0° and 24.6°C.

The animal could move around on the thin rubber platform of the recording unit illustrated in Figure 5. Compare these records with the kymograms shown in the next figure.

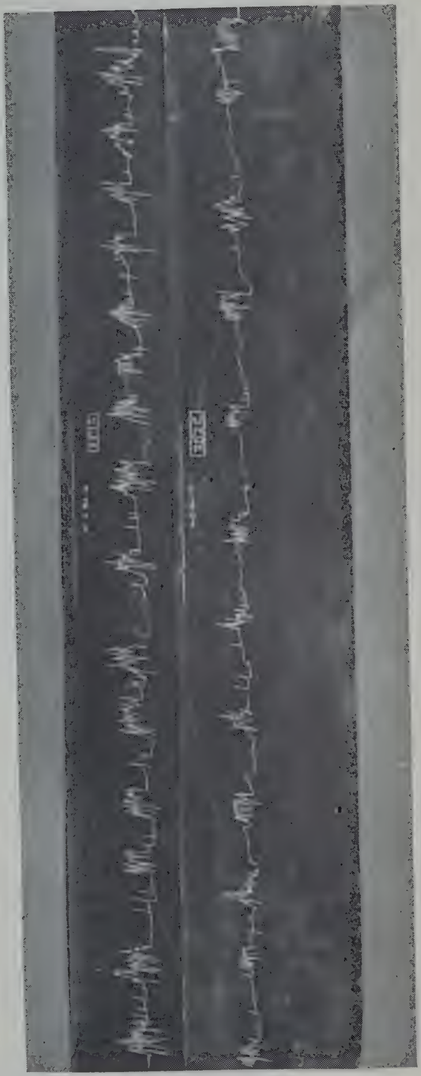


FIGURE 3

Records of Spontaneous Activity of a Young Mouse Held Upright on the Rubber Platform of the Apparatus Shown in Figure 5

Groups of movements now appear with great regularity.

Figure 3 contains records of the spontaneous activity of two-day-old mice which were prevented from tumbling over on their sides or backs. The animals were, in these cases, held upright astride a celluloid frame by a bit of adhesive tape attached to the back. The feet were always in contact with the thin rubber diaphragm of the recording apparatus (*cf.* Figures 5 and 6). At constant temperature the periods of movement are of about equal length; the probable error as percentage of the mean varies in comparable experiments from 2.5 to 8.1% (mean = 4.7%). The periods of quiescence also recur at regular intervals, but with greater variation (*cf.* Table 1).

TABLE 1

DURATION OF PERIODS OF ACTIVITY AND QUIESCENCE FOR DIFFERENT ANIMALS OF THE SAME AGE AND AT THE SAME BODY TEMPERATURE (21.0°C.)

		Duration of period of activity		Duration of period of quiescence	
Date	Animal No.	Av. (min.)	*P.E.*% of M.	Av. (min.)	*P.E.*% of M.
Jan. 28	1	0.61	±2.54	1.69	± 4.9
26	2	0.44	±4.20	1.56	±15.3
26	1	0.58	±3.20	1.75	±12.5
28	2	0.47	±8.10	1.95	±18.1
23	2	0.52	±5.30	1.72	± 6.4

We might expect one of two results when we record the activity of these young mice at different temperatures: (*a*) If the animal "willed" to move about, we would expect that the number of attempts to move per hour would not change with any simple regularity as we changed the temperature. (*b*) If, however, the organism acts as a physico-chemical whole, we would expect the frequency of the periods of activity to increase some 2 to 3 fold as we increase the temperature through 10°C. The data actually show an increase of 3.2 fold. This "temperature coefficient" (the well-known van't Hoff Rule) is only a crude empirical expression which gives us a ratio of rates at temperatures ten degrees apart, serving roughly as an index but having no theoretical significance. A better formulation of the effect of temperature is an equation derived empirically by Arrhenius and shown by Hinshelwood (13), Tolman (28), and others, to have theoretical foundations. It can be written

$$\frac{K_2}{K_1} = e^{\frac{\mu}{R}\left(\frac{1}{T_1} - \frac{1}{T_2}\right)},$$

where K_1 and K_2 are velocity constants (or figures proportional thereto) at the absolute temperatures T_1 and T_2; e is the base of natural logarithms, R is the gas constant, and μ, called by Crozier (7) the "temperature characteristic," is a constant with the dimensions of calories per gram molecule. It follows from the nature of the Arrhenius equation that when one plots the logarithm of frequency of periods of movement (assuming that these frequencies are proportional to the logarithm of the velocity constant of an irreversible chemical reaction at each temperature) against the reciprocal of the absolute temperature, a linear relationship is obtained from the slope of which the value of μ can be calculated. This was found to be 25,300 calories for these experiments (*cf.* Figure 10). From this evidence I feel certain that in these young mice the frequency of onset of the "impulse to move" behaves in a predictable manner obeying the laws of chemical dynamics.

II

It is necessary to go into the steps taken to estimate the bodily temperatures of the mice used in the experiments on spontaneous movements.

Working with young mice of another strain, Stier and Pincus (26) showed that when the outer temperature is changed the body temperature alters to a definite level above that of the environment and is then maintained constant at this new level. For two-day-old mice this equilibrium temperature was found to be 0.1°C. higher than the external temperature of 16°C., and from 2° to 3° higher at external temperature 34°C. The increase in the equilibrium temperature with increase in external temperature may be the resultant effect of the heat produced during the increased "spontaneous" muscular activity, or of general "basal metabolism" and of the rate of cooling which probably follows Newton's law of cooling (as found, for example, in chickens and in house wrens); cf. Mitchell and Haines (19), Kendeigh and Baldwin (16).

The differences between the internal temperatures of these mice and that of their environment were obtained by the method of thermometry commonly employed for such measurements (*cf., e.g.,* Rogers and Lewis, 1916; Pirsch, 1923). A primary thermopile, consisting of three insulated pairs of No. 40 copper-constantin junctions, was inserted into the rectum of the young mouse; the reference thermopile, consisting of three of the same bimetallic junctions joined in series with their corresponding primary junctions, extended 3 cm. into the beaker surrounding the mouse. A difference in temperature between the two thermopiles resulted in a deflection of the galvanometer. By direct calibration of this circuit, each mm. on the galvanometer scale represented a temperature difference of about 0.18°C. The temperature of the chamber containing the mouse was kept constant to ±0.05°C. or better by an outside water thermostat and by the circulation of air which was first bubbled through water and then passed through 20 feet of coiled glass tubing immersed in the water bath. A standard thermometer reading to 0.01°C., the thermopile leads, and the air line, entered the chamber through a hole in the plate glass window covering the beaker.

In determining the body temperature of the young mice employed for the experiments on spontaneous activity it was thought inadvisable to insert thermopiles into the rectum, since Pearcy and Hayden (20) had shown that stimulation from the region of the rectum may cause changes in irritability of the central nervous system. It is conceivable that the presence of a thermopile in the rectum might cause an amount of activity to appear in excess of that found for animals without a thermopile. I therefore determined separately the internal temperatures of 5 two-day-old albino mice of approximately the same mass as those employed in the experiments on spontaneous activity. These measurements were made between 14° and 33°C., by the method of inserting thermopiles into the rectum. From the data a graph (Figure 4) was constructed from which the body temperatures of the animals used in the experiments on activity could then be estimated. The points from different individuals deviate by negligible amounts from the average line. Such small discrepancies justify the use of this graph for approximating the body temperatures of the other mice of similar age and mass.

III

Two brass recording units, shown in Figure 5, were immersed in a water bath to a depth of two inches. Temperatures above that of the room were obtained by a bank of heating lamps, relay-controlled. "Sticking" and sparking at the contact surface of the mercury in the regulator was obviated by breaking a current of 0.2 microamperes which was then amplified through a triode valve so that it was sufficient to actuate a 5000-ohm telegraph relay; this system of thermoregulation has been described by Beaver and

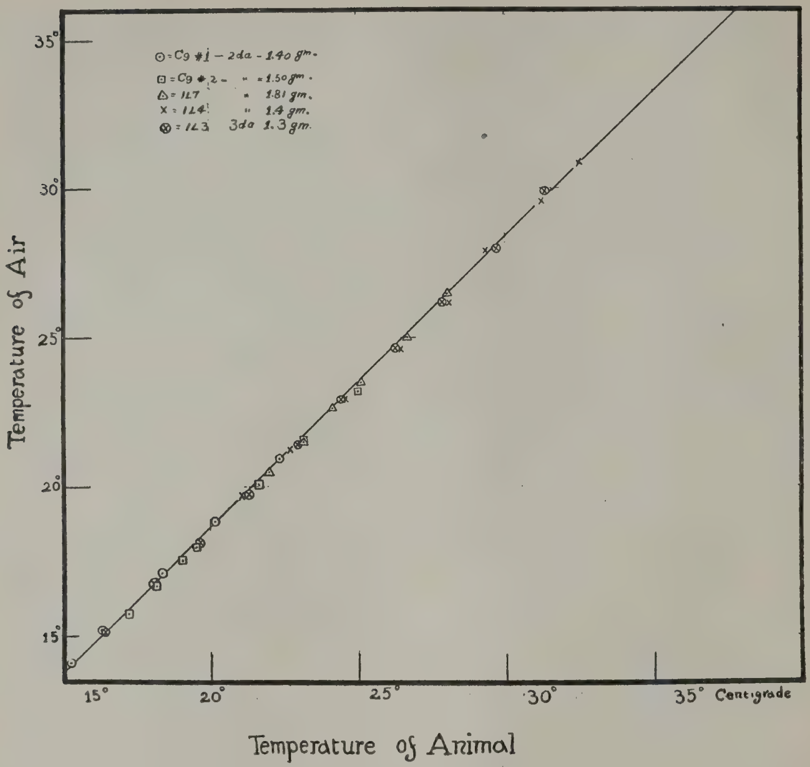


FIGURE 4

Plot Showing How the Body Temperature of a Two-Day-Old Mouse Varies with the Air Temperature

The temperature of the animal is always higher than that of the air. See text.

Beaver (1). With good stirring and sufficient felt insulation, the constancy of the temperature of the bath could be easily maintained to within ±0.005°C. The air in the brass chamber remained constant to within ±0.05°C. Fresh air from outside of the laboratory building was first bubbled through water and was then pumped through 40 feet of copper tubing (1/16-inch inside bore) into the brass chamber.

The movements of the animals were recorded upon glazed paper lightly smoked with soot from a kerosene lamp; this gives a better

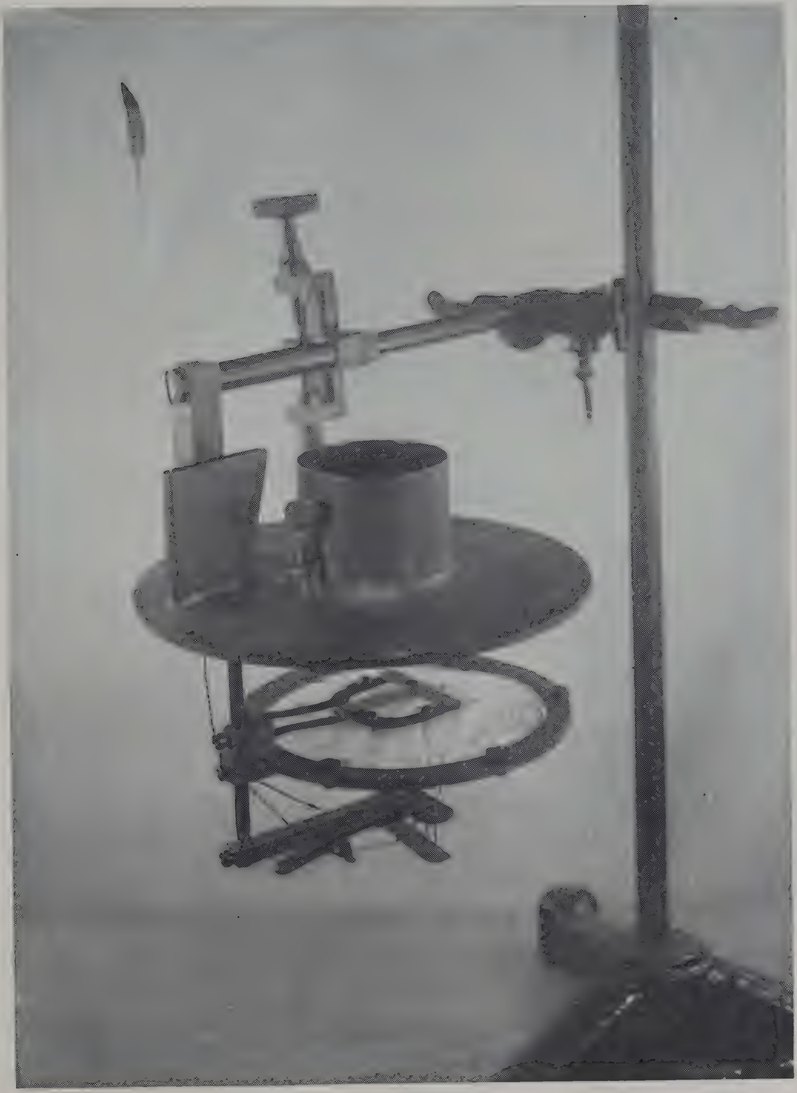


FIGURE 5

Apparatus Used for Recording the Movements of Young Mice

The animals are held upright astride a celluloid frame (marked *F*) by a strip of adhesive tape attached to their backs. This unit fits into a brass container (*cf.* Figure 6), a water-tight union being made by a rubber gasket. The container and recording unit can then be immersed in a water thermostat.

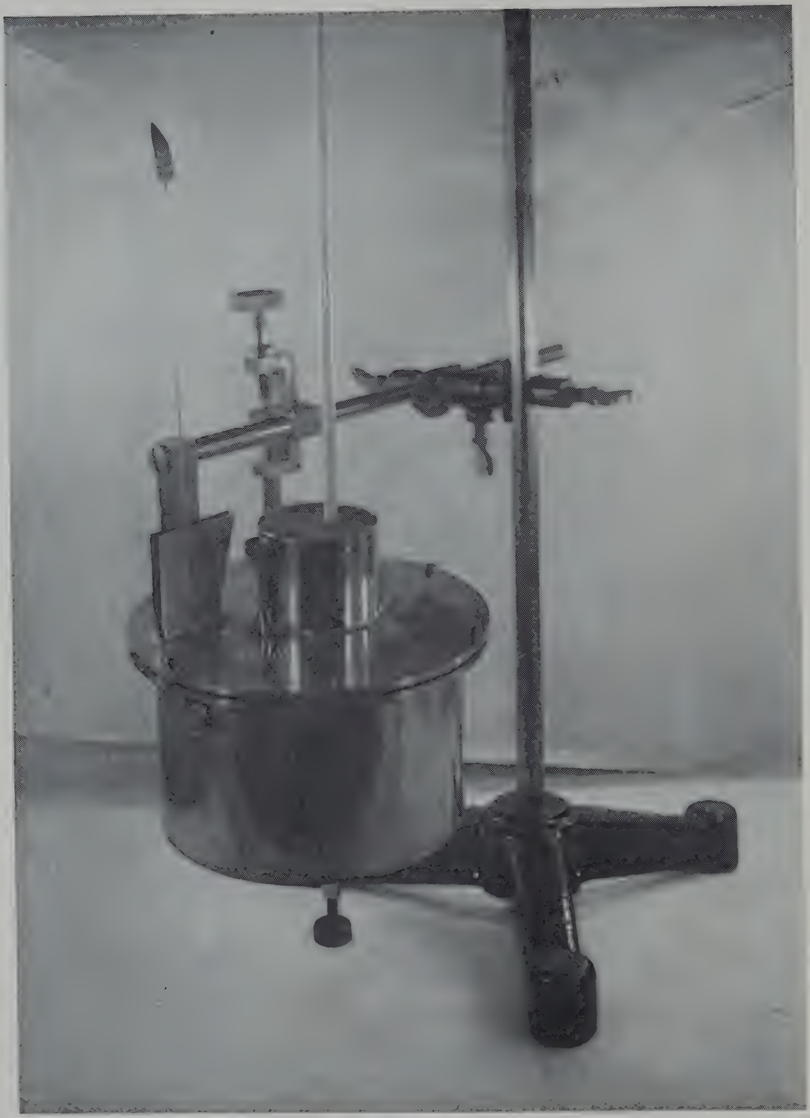


FIGURE 6

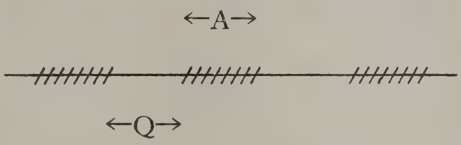
A PHOTOGRAPH OF THE RECORDING UNIT AS IT APPEARS WHEN ASSEMBLED AND READY FOR IMMERSION IN THE WATER THERMOSTAT

The various adjusting screws are used for delicately moving the writing point of the recording lever against the smoked paper of the kymograph. Air was admitted through an inlet placed in the bottom of the container (not shown in photograph).

recording surface than that usually prepared with a gas flame. A long-paper kymograph (Sherrington-Starling model), whose recording drums had been rotated through 90° to a horizontal position, was supported over the thermostat so that the writing lever of the recording unit made excursions in a plane parallel with the face of the kymograph drum. An electric governor kept the speed of the driving motor constant to within ±0.05%.

IV

The term "spontaneous movements" has been already defined to mean the bodily movements of the whole organism occurring when all the possible variables in the environment are held constant. I will use the term "spontaneous movements" synonomously with the expression "spontaneous activity." The meaning of expressions such as "period (or duration) of activity," "period (or duration) of quiescence," "cycle of activity," is indicated in the following diagramatic representation of a typical kymogram (compare with Figure 3).



A = Period of Activity
Q = Period of Quiescence
$Q + A$ = Cycle of Activity

The lengths of periods of activity and of quiescence on each kymogram were measured in inches and then converted into minutes by multiplying by a factor calculated for each kymogram from the length between marks made every minute by a signal magnet actuated through a clock mechanism. At least 20 cycles of activity were registered at each temperature. In all, 1070 periods of activity and 1010 periods of quiescence were obtained at 37 different temperatures.

Data collected at a series of temperatures with a single individual are shown in Table 2.

TABLE 2

RELATION OF TEMPERATURE TO THE LENGTH OF PERIODS OF ACTIVITY OF A TWO-DAY-OLD MOUSE

(Feb. 4, No. 4)

Air temperature	Body temperature	Duration activity (min.)	Duration quiescence (min.)	Av. number periods activity per hour
17.35°	17.7°	0.43±0.02	2.02±0.16	24.4
19.76°	20.20°	0.64±0.02	1.91±0.12	23.6
21.50°	21.95°	0.65±0.01	1.23±0.08	31.9
30.45°	31.10°	0.33±0.01	1.05±0.13	43.8

The actual picture of activity at different temperatures as it is recorded on smoked paper in these experiments can be seen in Figure 3.

When the mean duration of a single period of activity and of a single period of quiescence is plotted against the reciprocal of the absolute temperature, it is found that the slopes of the two lines are different—the constant, μ, in the Arrhenius equation, giving roughly a value of 13,200 calories for activity and one of 30,900 calories for quiescence (*cf.* Figure 8). These values indicate that the duration of quiescence decreases much faster with increase of temperature than the duration of associated activity decreases. This is seen in Figure 7.

Plotting the logarithm of the frequency of occurrences of movement per hour against the reciprocal of the absolute temperature gives a value of $\mu = 25{,}300$, with a shift of the line downward at about 25.5°C. Pincus, in his studies on the relation of temperature to frequency of respiratory movements in this same strain of mice, has obtained two classes of temperature characteristics—one of $\mu = 14{,}000$ and the other of $\mu = 33{,}000$. He finds a critical temperature at 25°C., where a shift downward occurs in the continuity of the line relating frequency of respiratory movements to temperature. (These experiments have not yet been published.) Crozier (8) showed that at certain temperatures irregularities of several kinds appear in the relation between vital processes and temperature. It was ascertained that these critical points are usually found to be in the neighborhood of 4.5°, 9°, 15°, 20°, 25°, 27°, 30°C.

The wide scatter of the points in some of the plots is indicative of the large variation which may enter into experimentation of this

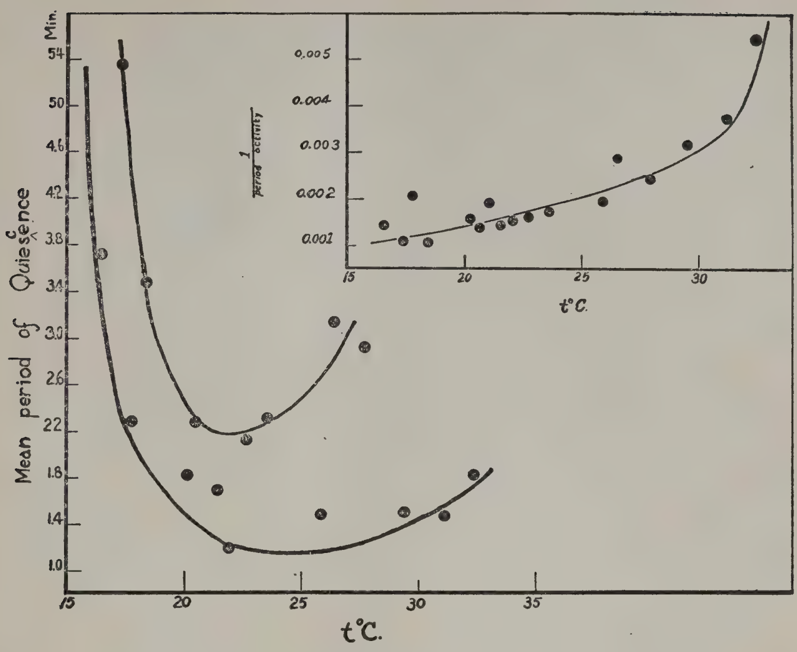


FIGURE 7

The Mean Duration of a Period of Quiescence is Plotted Against °C., and in the Same Figure the Reciprocal of Duration of an Activity Period against °C.

In the next figure these same observations are plotted according to the Arrhenius equation, *viz.* log 100/mean activity period, and log 100/mean period of quiescence, are plotted against the reciprocal of the absolute temperature. Each point is an average of data for 3 to 4 animals at a mean of 3 to 4 temperatures the extreme variation from the mean never being more than ±1°.

kind if all conditions are not adequately controlled. So far, in this investigation it has been impossible to use animals which were in exactly comparable states of development; or in entirely uniform states of nourishment. There are differences in the stages of development of the recently-born mouse. The period of gestation is sometimes shortened; it may even be lengthened by 2 to 8 days beyond the average. For future experiments it would be better to employ older animals, perhaps using the first appearance of pigmentation on the dorsum as a marker for obtaining animals which are in equal stages of development. Also, all nursing mothers do not feed their

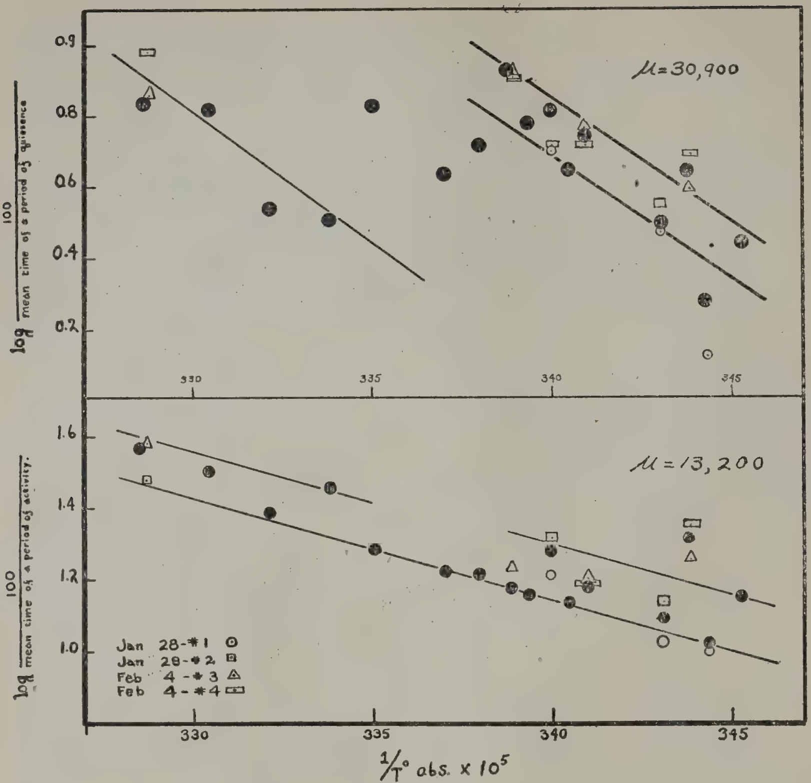


FIGURE 8

young with uniform regularity, or for equal lengths of time. This has a profound effect on the picture of spontaneous activity which is obtained. It was found important to remove the father from the cage a few days previous to the birth of the young, keeping him away during the course of the experiment, since it was noticed that there was greater negligence of the young by the mother when the male was present. This negligence may be due, in part, to the well-known fact that the female is impregnated by the male during the 24 hours after parturition. During this investigation about 50% of all litters were lost through eating of the young by the mothers.

Another source of variation in the plots is the lack of long series of determinations at different temperatures. Since one hour was allowed for adaptation at each new temperature, and since it was thought advisable not to use an animal for longer than 6 to 8 hours,

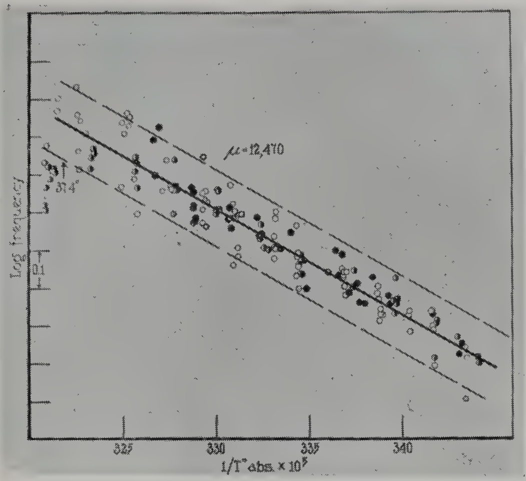


FIGURE 9

Mass Plot of Data for Three Mice Relating Log Frequency of Respiratory Movements to the Reciprocal of the Absolute Temperature
[Reproduced from a paper by Stier and Pincus (26).]

it was possible to obtain determinations only at three different temperatures for each pair of animals; six observations would have been more desirable. In Figure 9 a plot is reproduced in which temperature is related to frequency of respiratory movements of young mice of the piebald strain. However, the latitude of variation is as large in this plot as it is in Figure 10, where temperature is related to frequency of occurrence of periods of spontaneous movements. The variation in performance is thus largely "intrinsic," but not lawless.

In some earlier experiments with a strain of piebald mice, in which it was attempted to relate temperature to the frequency of occurrence of righting movements (that is, attempts of the animal to raise itself when it was held in a harness, dorsum down), I obtained a value of $\mu = 22{,}500$ with a "break" at 20°C. The data are few, so I present this value merely as being suggestive. It is of interest that in our studies of the relation of temperature to the respiratory movements in this same strain of mice, we obtained a critical temperature at 19.8°C. [Stier and Pincus (26)]. Observations by Tracy (29) on the development of spontaneous movements of embryonic toadfish are of interest in this connection. He suggested that movements of the whole embryo in these early stages may serve as a kind of respiratory activity. His experiments show that a condition of asphyxia obtained by producing a high CO_2 concentration in the environment affects simultaneously both spontaneous

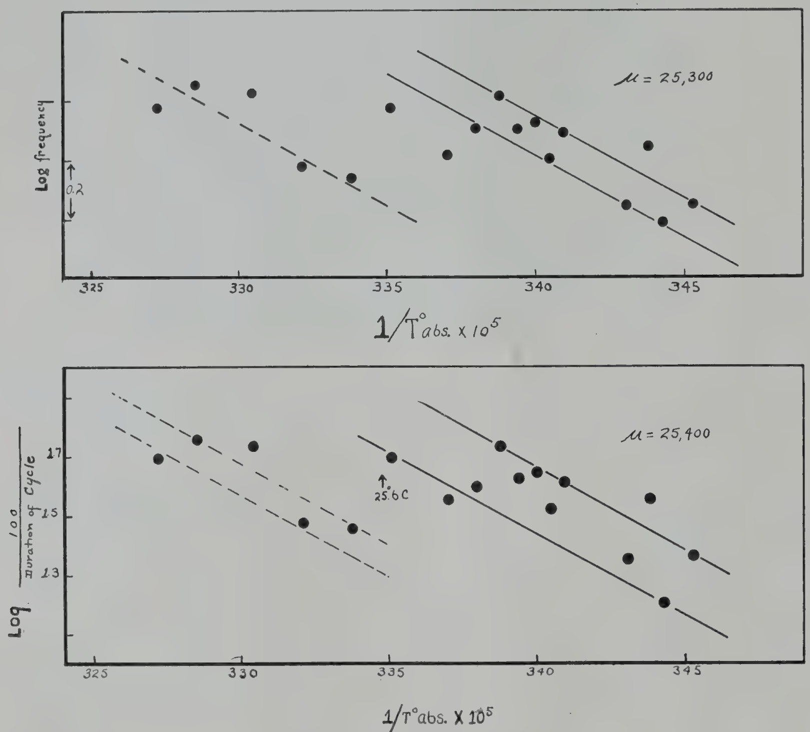


FIGURE 10

Relation of Temperature to the Number of Activity Periods per Hour in Two-Day-Old Mice

Notice the "critical temperature" which appears at 25°; such a shift downward of the line relating *temperature to log frequency* is also obtained by Pincus in this strain of mice for the frequency of respiratory movements as related to temperature. The fit of the lines to the data was based partly on this information and partly on numerous experiments with a variety of animals where similar "shifts" sometimes occur in plots where temperature is related to frequency of heart-beat and respiratory movements [Crozier and Stier (11)]. Without this information at hand a line parallel with the X axis might have been fitted to the observations. The heavy circles refer to averages of data for 3-4 individuals. Actual values of the observations can be found in Figure 7.

movements and respiratory movements equally. Experiments by Brown (1915), on movements of foetal cats thought to be produced by asphyxia when the umbilical cord is compressed, suggested the possibility of this relationship. The close similarity of critical temperatures found for both spontaneous movements and respiratory movements in the albino and piebald strains of mice, but differing in the two strains, may perhaps be suggestive of some kind of a relation between the respiratory center and the mechanism controlling spontaneous activity.

From these data it can be concluded that there may be two separate processes which independently control the component parts of a cycle of spontaneous movements: one process determining the duration of quiescence, the other controlling the span of activity. For a working hypothesis it is assumed that the following mechanism controls the initiation of activity: a substance A is built up from inactive substance A'. When A reaches a threshold concentration, vigorous movements begin; when A is reduced below the threshold amount, activity ceases. As the temperature is increased, the rate of formation of A is speeded up, thus causing an earlier appearance of activity. (*Cf.* Figure 7 and Figure 8, where it is seen that as the temperature increases the length of the period of quiescence shortens.) This *kind* of mechanism has been found to control the duration of a state of tonic immobility, usually called "death feigning" [Crozier (6), Crozier and Federighi (9)]. The duration of the immobile state in arthropods was assumed by Crozier to be proportional to the amount of a substance B formed in a linked chain of chemical reactions as follows: $A \rightarrow B \rightarrow C$. Hoagland (14) found that for the lizard *Anolis carolinensis* the duration was also perhaps determined by an autocoid (possibly adrenin) released in definite amounts by the stimulation used to induce tonic immobility.

In a later section, several suggestions are made as to the possible nature of the substance A, and the probable locus of its action. I must first, however, present additional experiments on the spontaneous activity of the same albino strain of mice as I have employed thus far, but at the ages of 40 and of 90 days.

"Spontaneous Activity" in Older Mice

I

A photograph of the apparatus devised for recording the activity of these older mice is given in Figure 11. It is patterned somewhat

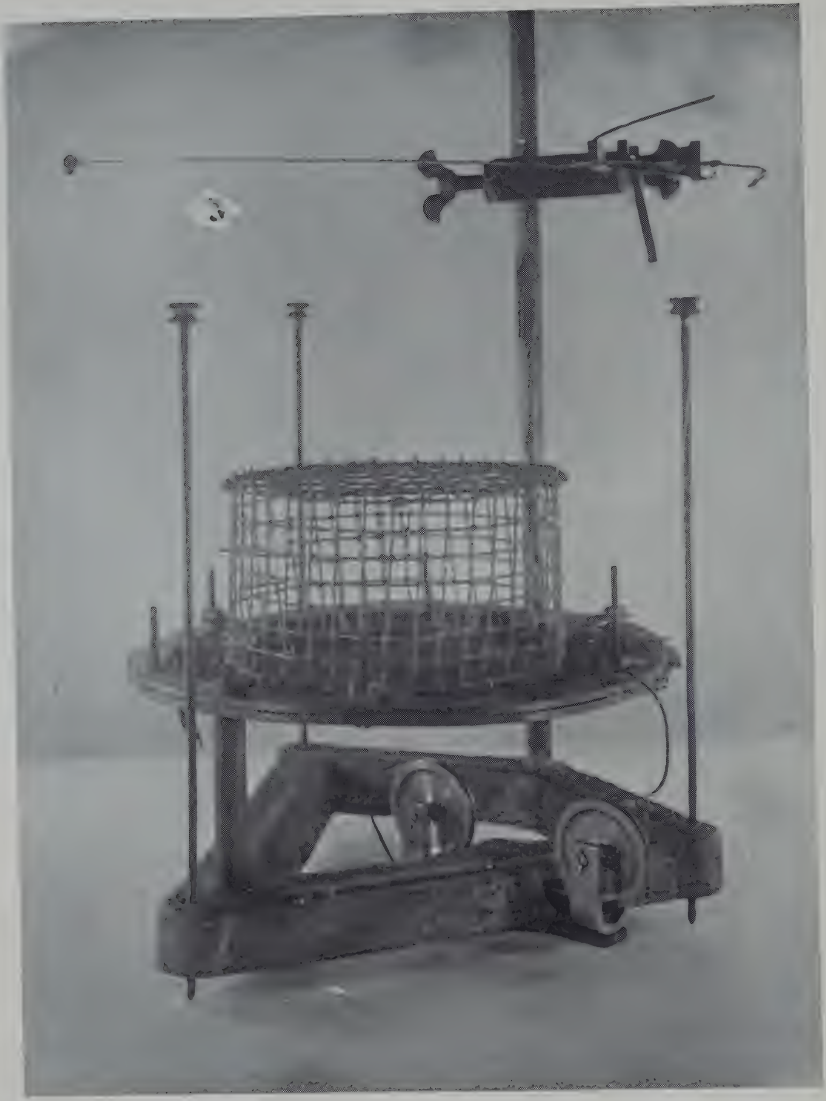


FIGURE 11

Apparatus for Recording the Activity of Adult Mice

A thread cemented to the center of the rubber dam platform transfers the excursions of the platform to a heart lever by way of the two pulleys. The wire cage rests on the rubber diaphragm so that any climbing movements would be recorded. This apparatus registered displacements of as small a mass as one-half gram. Water was always available in the small glass container fastened to the side of the cage.

after a device used by Richter (23) in his experiments on the activity of young rats. The recording unit (Figure 11) was lowered into a large tin can. Three such units were mounted on frame work, whose supporting braces were securely bolted to a thick brick wall, insuring freedom from vibration, and which held the cans, to within an inch of their tops, in a large water thermostat. The temperature of the bath was kept constant at $23.0^\circ \pm 0.05^\circ$C. by adding heat from two banks of electric lamps by a simple thermoregulation system consisting of a large mercury thermoregulator and a 150-ohm telegraph relay. "Chattering" of the relay was entirely eliminated by having the contact needle in the thermoregulator of copper instead of nichrome. Even after four weeks of constant use little of the copper had been lost through amalgamation. A small constant stream of air from out-of-doors was first bubbled through water (to insure constant humidity at all times) and then forced into the three recording units by way of 60 feet of copper tubing immersed in the water bath.

The experiments were conducted in a dark room which was practically sound proof. (Tests in which loud noises were purposely made resulted in no responses from quiet mice.) The intensity of the light in the dark room was kept constant day and night throughout the course of the experiments. The containers of the recording units were covered with a sheet of isinglass so that their interiors were at about the same light intensity as the cages in which the experimental animals lived.

Mice of the same albino strain, and of the same genetic composition as those used for the researches previously discussed, were employed for these experiments. A definite feeding schedule was started a week before the actual recording of activity began and was maintained throughout the entire period of experimentation. This consisted of one feeding per day. Each day at 11:00 A.M. dishes filled with food were placed in the cages, were left there for 1½ hours, and were then removed. Water in glass containers was always available to the animal. When lettuce is supplied twice a week, and the cages are kept clean, animals can exist on such a meager feeding schedule. Deaths, however, were frequent. The 90-day-old animals seemed to survive much better than their younger relatives.

During the entire investigation the animals were kept in the same dark room where the experiments were conducted. This was done with the intention of "adapting" the animals to the noises of the

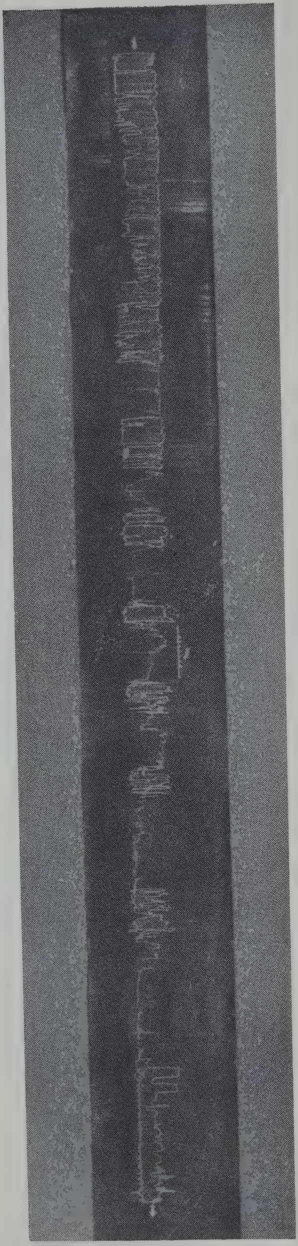


FIGURE 12

This record was obtained when a forty-day-old mouse was placed in the apparatus shown in Figure 11. The activity was recorded from 2:00 P.M. until noon of the following day. The time marks are five minutes apart. Periods of activity were outlined with white ink in order to make them stand out more clearly.

electric motors which were used. The volume and quality of the noise remained practically constant during the whole period of these experiments.

The records of the spontaneous movements were obtained on smoked, glazed paper by a method already described. A clock mechanism fitted to actuate a signal magnet every 5 minutes marked time-intervals on the kymogram. Immediately upon removing the food from the cages the animals were placed in the recording apparatus and the kymograph was started.

II

In Figure 12 a record of the spontaneous activity over a period of 22 hours is presented. There is a great similarity to the kymograms obtained for movement of two-day-old mice. At this age, however, the picture is somewhat changed. There are still periods of great movement followed by periods of quiet; the periods of activity are still initiated by the same explosiveness and end with the same abruptness:

Two-day-old mice—duration of activity 0.3 to 0.6 minutes
duration of quiescence 1 to 3 minutes
Forty-day-old mice—duration of activity 25 to 135 minutes
duration of quiescence 15 to 160 minutes

These older mice are active for longer periods of time and are also quiet for longer periods of time.

The kymograms which were obtained as records of the activity of these mature mice can be fairly classified into two groups.

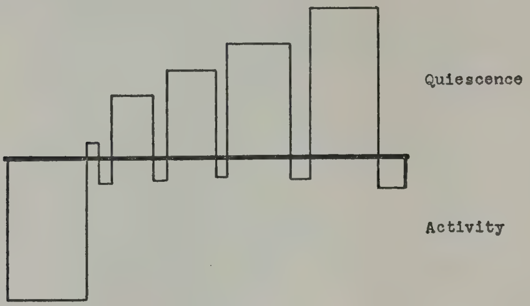
Example of Class A Group
(See text and figure on following page)

Total activity	Total quiescence	A/Q
665.0 min.x5	485.5 min.x5	1.4
572.0*	668.5	0.9
422.0*	577.5†	(0.7)
633.5*	595.0†	1.0
610.5	490.0	1.23
263.5*	649.5†	0.41
533.5*	687.5†	0.80
545.0	769.0†	0.70
404.0*	691.5†	(0.58)
515.5	511.5	1.30
735.5	536.0	1.40

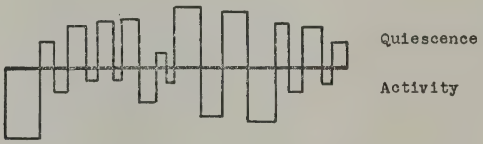
*typical initial outburst of activity
†typical progressive increase in quiescence
Other experiments belong to Class A but are less typical.

Class A: An initial outburst of activity lasting from feeding time until 2:00 P.M. (3 hours), followed by short periods of activity which were accompanied by progressively lengthening periods of quiescence.

Fourteen runs on 3 comparable individuals gave 9 series in Class A and 5 in Class B.



Class B: No initial outburst of activity. The periods of activity and of relative quiescence are of varying length; the sum of both being almost equal.



EXAMPLE OF CLASS B GROUP

Total activity	Total quiescence	A/Q
623.5 min.x5	625.5 min.x5	1.0
588.5	591.5	1.0
661.0	669.5	1.0
601.0	546.5	1.1
(212.5)	(150.5)	(1.4)
537.5	556.0	1.0
596.0	562.0	1.1

An examination of the accompanying tables will show the actuality of these two classes. Class B seems to be the more stable, giving a ratio of activity/quiescence equal to 1.0 as the mode of all the cases. The total activity in 12 hours is almost constant for the different animals on different days.[1]

Class A, although it shows definite form in the kymograms (*cf.* Figure 12), has values for the component parts of its cycle which show wide variation. More detailed investigation will perhaps reveal the nature of these two classes.

Part of the program of the present investigation was to find out if there is some difference in the picture of activity of two strains of mice, known to be respectively homozygous for the specific characters *albino* and *dilute brown.* It is planned to study the inheritance of the particular characteristics of activity for each race when the two are crossed.

Casual observation had shown that mice of the dilute brown race are very "nervous"; the animals jump around with great vigor, and are more difficult to handle than the albino race. The latter individuals are more docile and appear to be slower in their movements. My measurements show a difference in the total duration of activity per day in the two races. The albinos are active for longer periods of time than the dilute browns. These data, however, only give duration of activity in minutes. Attempts to develop an apparatus for estimating the amount of work done in each period of activity have been unsuccessful thus far. Future experiments will attempt to estimate this quantity. It will then be possible to describe more accurately these two races in quantitative terms rather than in descriptive terms, saying merely that one is more lively than the other.

There is a definite change with age in the total duration of bodily movements, the animals becoming active for longer periods of time as age advances.

This change with age has been well established for rats by Richter (23) and by Slonaker (25). These investigators found that the amount of activity (number of marks on the kymogram, Richter; number of revolutions of a revolving cage, Richter and Slonaker) increased from birth, reaching a maximum at 100 to 175 days of

---

[1] The regular occurrence of two peaks of activity each day at about 7 o'clock (sundown) and at 5 o'clock (sunrise) has been found in many animals by Szymanski (27), and has been carefully observed in the Japanese waltzing mouse by Wolf (33). The experiments of Richter (22) on white rats and my own records do not show this regular increase of activity. Perhaps the experimental procedure of feeding only once each day has modified this cycle.

TABLE 3

SUM OF THE DURATION OF ALL PERIODS OF ACTIVITY DURING 12 HOURS FOR TWO RACES OF MICE, ALBINO AND DILUTE BROWN
Six females, each 40 days of age.

	Albino (3)	Dilute brown (3)
	1755 min.	1220 min.
	2080	1200
	1795	1110
	1790	1050
	1660	
	1525	
	1855	
	1990	
	1885	
	1855	
	1040	
	1525	
	1495	
Average	1715±49.3	1145±31.7

TABLE 4

CHANGE IN THE DURATION OF ACTIVITY PER 12 HOURS WITH AGE

	40 days of age	90 days of age
Albino	1715 min.	2060 min.
Dilute brown	1420	2030

age. It then decreased gradually until death. This maximum of movement almost coincides with the end of maximum fertility as seen in the curve of number of litters cast in the white rat (King, 1916). This points to the influence of a sex hormone on the spontaneous activity of the rat.

The influence of some sex hormone on spontaneous activity can be seen in the distribution curve which is obtained when the total duration of activity per 12 hours for all albino females is plotted (Figure 13). The curve is distinctly bimodal, the modes occurring at 1450 and at 1750 minutes. In light of the experiments on the relation of the oestrus cycle to activity in rats, one might expect such a bimodality to occur. Many careful and extensive experiments [Wang (32)] show that in female rats there is an increase of total duration of activity occurring regularly every four days. This increase occurs simultaneously with ovulation. Injection of the ovarian hormone in spayed, or even in castrated rats, produces the same increase in activity. The exact nature of this increase, that is, the way the amount of work done increases in each cycle of activity, will be investigated later.

In Figures 14 and 15 there are plotted the distribution curves for the duration of each period of activity during each day that the animals were observed. The same is done for the duration of individual periods of quiescence. There are some curious relations in the multimodes which seem to be in evidence. Let it be said, however, that these relationships are only suggestive—not until many more experiments are performed can accurate deductions be drawn.

The frequency distribution of periods of quiescence (Figure 14) gives possible modes at 15, 25, 37.5, 55, 65, 82.5, 110, 140, and 160 minutes; the differences between these modes are 10, 12.5, 17.5, 10, 17.5, 27.5, 30, and 20 minutes. These differences give an average of 14.7 minutes. Thus, when a period of quiescence occurs, it is very nearly some multiple of 15 minutes plus a constant (15 minutes).

Inspection of the distribution of activity periods (Figure 15) shows that there is also more than one mode. These modes are not very distinct. They tend to come at . . . 25, 50, 65, 95, 115, and 145

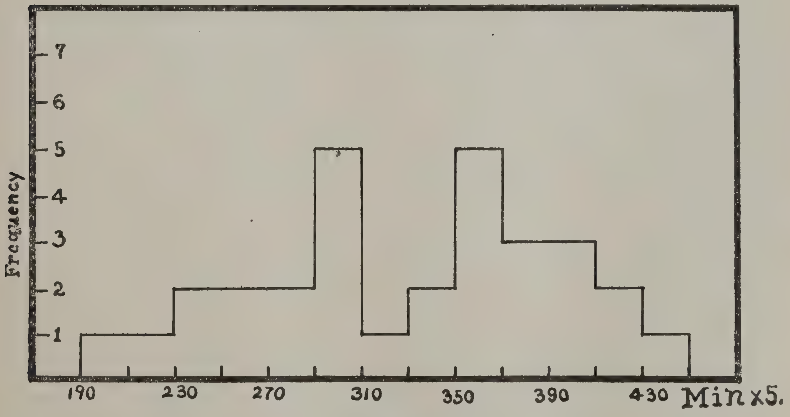


FIGURE 13

A Curve of Distribution of the Total Duration of Activity in Twelve Hours

See text for discussion.

minutes; the difference between these modes are 25, 15, 30, 20, and 30 minutes, and the average difference is 24 minutes.

Whenever a period of activity occurs it is likely to be for some multiple of 24 minutes plus a constant (25 minutes). When activity and quiescence increase they do so in steps which are nearly some multiple of their mode containing the shortest periods of time. The average difference between modes in both activity and quiescence is equivalent to the mode containing the shortest periods.

We will assume that there is a basic cycle of activity consisting of

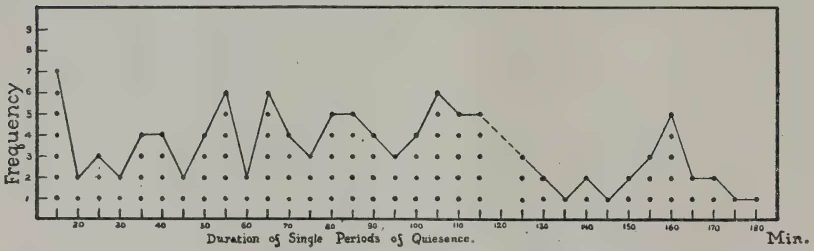


FIGURE 14

A Curve of Distribution for the Duration of Single Periods of Quiescence

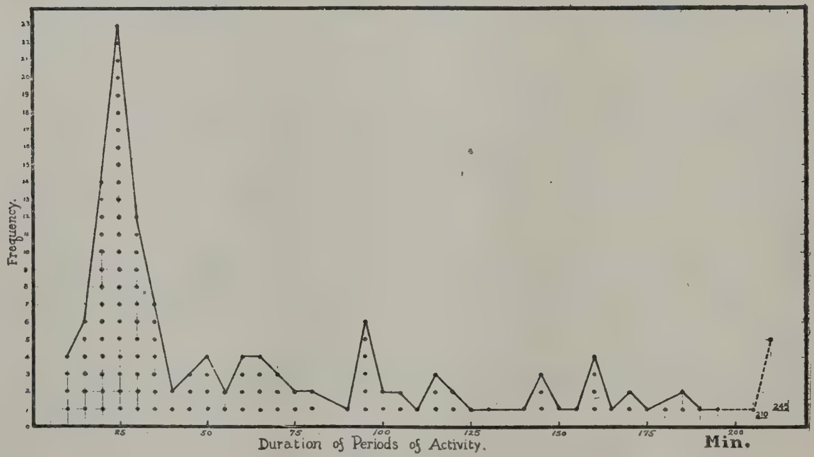


FIGURE 15

The duration of single periods of activity when placed in classes five minutes apart give a curve of distribution which is multimodal.

periods of quiescence (Q) and of activity (A) which follow each other with this time relation:

Q	A	Q	A	Q	A	Q	A	
15	25	15	25	15	25	15	25	minutes

If any variation occurs in this cycle, that is, periods of activity which lengthen or periods of quiescence which lengthen, it is found that the periods of quiescence become larger in average steps of 14 minutes; the periods of activity increase in multiples of 24 minutes.

This leads us to the assumption that the substance A, our postulated material for initiating groups of activity, can be delivered in doses or quanta. A similiar instance was found by Crozier and Federighi (9) and by Pincus (experiments unpublished), when they caused isopods to go into states of tonic immobility. They found that successive cycles of immobilization and recovery therefrom were sometimes some definite multiple of the lowest basic length.

III

It is not the purpose of this account to present a complete review of the numerous investigations which have been made on the problem of spontaneous activity. We will have to limit the discussion to a rapid history of one of the theories which seems quite well substantiated, *viz.,* that the periods of activity of the organism are highly correlated with the periods of "hunger movements" of the stomach.

This idea was proposed in essence by Jacques Loeb in 1900, although at that time nothing was known about the periodicity of the hunger contractions of the stomach. In explaining the curious behavior of *Nereis* when it becomes unceasingly active after it has been deprived of its supraoesophageal ganglion, Loeb writes, "The supraoesophageal ganglion of Annelids is connected with the alimentary canal by nerves. The processes which go on in the intestine . . can only affect the whole animal nervously as long as the supra-oesophageal ganglion is intact. If it is removed the whole influence of this so-called sympathetic nervous system ceases. It is possible that the stimuli which pass from the sympathetic into the central nervous system may condition the alternation of rest and activity which the normal animal exhibits, and that the removal of this stimulation may remove the necessity of resting" (17, p. 95). He also precluded the possibilities of a humoral control.

Without apparent knowledge of Loeb's ideas about spontaneous activity, or of Boldyreff's findings about periodicity in organisms, Carlson reached these same general conclusions (5, Chap. I). Carlson primarily concerned himself with the nerve impulses which control the hunger mechanism and not with the mechanism governing the movements of animals.

Whether the Pavlov group of workers was directly influenced by Loeb's idea is not known at the moment. In 1904, one of Pavlov's students, W. Boldyreff, started to record the contractions of the intact stomach of dogs and to investigate the changes which occur simultaneously in other organs. His insight into the influence of the contractions of the empty stomach on the total bodily movements of the organism seems to have been immediate. His interest turned from the nature of the "hunger contractions" of the stomach to the nature of the periodicity of organic activity in mammals. From his researches [*cf.* Boldyreff (2)] one learns that there are periodic fluctuations in the activity of the various glands and organs in an animal. Curiously enough these fluctuations all occur simultaneously: when the empty stomach starts contracting, the pancreas, the small intestine, and the gall bladder start secreting; at the same time various enzymes appear in recognizable concentrations in the blood. Boldyreff was the first to show that the "spontaneous activity" of an animal is highly correlated with the vigorous contractions of the empty stomach. He found that dogs are restless and move about during periods when the empty stomach contracts most actively [so-called "hunger contractions," *cf.* Carlson (5)]; and that dogs stand quietly during the times when the stomach is relatively quiet. No mention seems to have been made by him about the elements which work together to control such periodic functions in the organism. A hint as to some humoral correlating mechanism was given, but no importance was attached to any afferent impulses which might be coming into the central nervous system from the gastro-intestinal tract (5).

Finally, in 1922, experiments were published which might be considered as tests of Loeb's original suggestion. Rather extensive experiments were made by Miss Tomi Wada (31) on human subjects, recording at the same time hunger contractions and bodily movements during sleep. The experiments are varied, but the measurements are few. She found that the bodily movements of the sleeping person occur only during the most intensive movements of the empty stomach, the so-called "hunger contractions" [Cannon and Washburn (4), Carlson (5)]. Such contractions can be initiated in the "quiet stomach" by increasing the intragastric pressure, whereupon movement of the animal results. Food taken before retiring delays the typical sequence of hunger contractions of the relatively empty stomach and delays the occurrence of the subsequent spontaneous movements. Drugs which relax smooth muscle, so-called spasm-inhibiting drugs such as papaverin, offset the periodic occurrence of hunger contractions and also the bodily movements of the organism.

In the same year (1922) Richter (22) published intensive experiments on the spontaneous activity of rats of different ages and under numerous experimental procedures. He found that periods of activity recur in rats about every hour and last approximately one-half hour when the animal is fed once a day. Whereas when animals are allowed to eat freely, the periods of activity occur every two hours and last about one hour [*cf.* also Richter (23)].

IV

From a consideration of the data obtained in experiments designed to relate spontaneous activity to temperature, it was proposed that a material *A* might be formed whose presence above a certain threshold concentration might be regarded as necessary for initiating and maintaining locomotor activity. From the evidence presented in the preceding section, it was concluded that movements of an intact animal are correlated with and determined by the vigorous "hunger contractions" of the empty stomach. This leads us to the assumption that the material *A* initiates the periodic recurrence of heightened tonus and of vigorous contractions in the stomach. The movements of an animal are then brought about through afferent impulses originating at the time of the vigorous stomach contractions. These impulses upon entering the medulla oblongata [*cf.* Carlson (5)] might be thought to cause a temporary cessation of the inhibitory action of the cerebral cortex over the lower centers. An animal would then behave like a *Nereis* whose supraoesophageal ganglion had been removed—it would be active [Loeb (17)]. Fulton (12, p. 526) has reviewed instances where the cerebrum and centers adjacent to it act as regions of origin of inhibition.

To adherents of a "humoral" theory of inhibition [*cf.* Fulton (12)] the notion might appeal that the material *A*, whose locus of formation, it has been proposed, is in the alimentary tract, may act in reducing the effective concentration of an inhibitory substance whose antagonistic action on a specific excitatory substance is thought to take place within or about the perikarya of nerve cells; or the substance *A* might be thought to alter synaptic barriers so that the central pathways of progression reflexes might be readily completed. That a hormone is present in the blood stream during the vigorous "hunger contractions" is evidenced by the experiments of Ivy (15) on the nature of the humoral control of the transplanted stomach pouch, the contraction of an isolated segment of intestine, the secretion of a transplanted piece of pancreas. Experiments by Boldyreff (2), by Sineinikoff (24), and especially blood transfusion experiments by Tschukitscheff (30), corroborate the notion that the periodic motor activity of the empty stomach is controlled by hormones which periodically rise to a maximum concentration in the blood stream. The locus of these autocoids has been shown by W. N. and A. W. Boldyreff (3) to be the pancreas. Experiments are in progress to test these notions of the way in which substance *A* could initiate activity. A summarizing statement of the mechanism proposed for the control of spontaneous activity follows.

It seems to me that the facts from existing investigations and from these experiments can be understood on the following basis: A material, call it *A*, is discharged rhythmically or periodically in doses of a certain magnitude from some organ serving as a "center"; the frequency with which this discharge occurs could be determined by the velocity with which some underlying process causes the material *A* to accumulate in the loci from which it is subsequently discharged —the frequency of these discharges would therefore bear a definite relation to temperature if at each discharge there occurred an outburst of "activity"; it could then be presumed that the activity would continue until the concentration of the effective substance *A* is reduced to a lower threshold; this must be presumed because the activity does disappear and then recur. The disappearance and recurrence cannot be understood in terms of exhaustion and recovery within the central system because the periods of activity are directly correlated with the time of independent outburst of motility in the alimentary tract; the degradation of the material *A* might be presumed due either to its dissemination throughout the body (that is, to its dilution), or, on the other hand, to its *chemical* removal; in the latter case the reciprocal of the duration of a mean period of activity should exhibit a characteristic high critical increment ($\mu =$ 13,200 was found, *cf.* Figure 8); in general it would be expected that the temperature characteristic for the frequency of outburst of activity and for the duration of activity would not be the same (*cf.* Figures 8 and 10). $\mu = 25{,}300$ was found for frequency of activity cycles, $\mu = 13{,}200$ for duration of periods of activity. The dilution might be due to the distribution of material through the circulation, and hence a function of the speed of circulation and thus of the frequency of the heart-beat; in general, heart-beat frequency and breathing rhythm are rather closely correlated; in these mice the μ for breathing frequency is 14,000 and 34,000. However, there is no necessity to assume that μ for breathing and heart-beat is the same.

SUMMARY

When every effort is made to eliminate excitation in the environment of a mouse, movements still occur—the so-called "spontaneous movements." These movements do not occur haphazardly, but with a definite form, namely, in regularly recurring surges of movement followed by periods of quiet; at constant temperature over intervals of observation lasting about two hours the periods of activity and of quiescence are of constant length, having a mean probable error, as percentage of the mean, of 5% for activity, of 10% for quiescence (at 21°C.). The mice are still poikilothermic at the age of two days; the frequency of activity cycles increases with temperature in the manner of a chemical reaction. When the average frequency of occurrence of activity periods is plotted according to the Arrhenius equation, the value of μ was found to be 25,300 calories. The duration of intervals of quiescence and of activity also changes as a function of temperature; when their reciprocals are plotted according to the Arrhenius equation, the value of $\mu = 13{,}200$ is obtained for activity, $\mu = 30{,}900$ for quiescence. It was suggested as a working hypothesis that these different values indicated that at least two processes were responsible for control of activity and of quiescence. It was assumed that activity results when the concentration of a substance *A* is maintained above a threshold level. Quiescence begins when *A* is reduced below this level, the velocity of the formation and of the removal being controlled by chemical changes.

Mice 40 to 90 days of age show the same cyclic occurrence of

periods of activity as exhibited by the two-day-old mice. Variation from an assumed basic cycle occurs in multiples of 15 minutes for activity, in steps of 24 minutes for periods of quiescence.

The mechanism controlling the "spontaneous" locomotor activity of mammals may have for its "center" the gastro-intestinal tract, since movements of the empty stomach are correlated with the incidence of bodily activity. Evidence collected from published accounts pertaining to spontaneous movements shows that movements of an intact animal are correlated with and determined by the vigorous "hunger contractions" of its empty stomach.

REFERENCES

1. BEAVER, D. J., & BEAVER, J. J. A new method for the control of thermostats. *Indus. & Eng. Chem.*, 1923, **15**, 359-366.
2. BOLDYREFF, W. Fonction périodique de l'organisme chez l'homme et les animaux d'ordre superieur. (Pancreas comme principal agent du processus de l'assimilation dans tout le corps.) *Quar. J. Exper. Physiol.*, 1916-17, **10**, 175-201.
3. BOLDYREFF, W. N., & BOLDYREFF, A. W. What is the cause or the mechanism of the periodical (hunger) gastric and intestinal contractions? *Skand. Arch. f. Physiol.*, 1926, **49**, 98-99
4. CANNON, W. B., & WASHBURN, A. L. An explanation of hunger. *Amer. J. Physiol.*, 1912, **29**, 441-454.
5. CARLSON, A. J. The control of hunger in health and disease. Chicago: Univ. Chicago Press, 1916. Pp. vii+319.
6. CROZIER, W. J. Reflex immobility and the central nervous system. *Proc. Soc. Exper. Biol. & Med.*, 1923, **21**, 55-56.
7. ————. On the possibility of identifying chemical processes in living matter. *Proc. Nat. Acad. Sci.*, 1924, **10**, 461-464.
8. ————. Note on the distribution of critical temperatures for biological processes. *J. Gen. Physiol.*, 1925-26, **9**, 525-529.
9. CROZIER, W. J., & FEDERIGHI, H. On the character of central nervous processes. *Proc. Soc. Exper. Biol. & Med.*, 1923, **21**, 56-57.
10. CROZIER, W. J., & PINCUS, G. Tropisms of mammals. *Proc. Nat. Acad. Sci.*, 1926, **12**, 612-616.
11. CROZIER, W. J., & STIER, T. J. B. Thermal increments for pulsation-frequency in "accessory hearts" of Notonecta. *J. Gen. Physiol.*, 1926-27, **10**, 479-500.
12. FULTON, J. F. Muscular contraction. Baltimore: Williams & Wilkins, 1926. Pp. xv+644.
13. HINSHELWOOD, C. M. The kinetics of chemical change in gaseous systems. London: Oxford Univ. Press, 1926. Pp. 204.
14. HOAGLAND, H. The mechanism of tonic immobility ("animal hypnosis"). *J. Gen. Psychol.*, 1928, **1**, 426-447.
15. IVY, A. C., *et al.* Contributions to the physiology of gastric secretion. *Amer. J. Physiol.*, 1925-28, **74-83**.

16. Kendeigh, S. C., & Baldwin, S. P. Development of temperature control in nestling house wrens. *Amer. Natur.*, 1928, **62**, 249-277.

17. Loeb, J. Comparative physiology of the brain. New York: Putnam, 1900.

18. MacDowell, E. C., & Lord, E. M. Data on the primary sex ratio in the mouse. *Anat Rec.*, 1925, **31**, 143-148.

19. Mitchell, H. H., & Haines, W. T. The critical temperature of the chicken. *J. Agric. Res.*, 1927, **34**, 549-557.

20. Pearcy, J. F., & Hayden, D. B. Reflexes from the gastro-intestinal tract to the labyrinth. *Amer. J. Physiol.*, 1928, **87**, 196-199.

21. Pearl, R. A time recorder for quantitative work in animal behavior. *J. Wash. Acad. Sci.*, 1927, **26**, 541-543.

22. Richter, C. P. A behavioristic study of the activity of the rat. *Comp. Psychol. Monog.*, 1922, **1**, No. 2. Pp. 55.

23. ————. Animal behavior and internal drives. *Quar. Rev. Biol.*, 1927, **2**, 307-343.

24. Sineinikoff, E. I. Über Hungerbewegungen eines Dünndarmabschnitts beim Menschen. *Pflüg. Arch. f. d. ges. Physiol.*, 1926-27, **215**, 427-430.

25. Slonaker, J. R. Long fluctuations in voluntary activity of the albino rat. *Amer. J. Physiol.*, 1926, **77**, 503-508.

26. Stier, T. J. B., & Pincus, G. Temperature characteristics for frequency of respiratory movements in young mammals. *J. Gen. Physiol.*, 1927-28, **11**, 349-356.

27. Szymanski, J. S. (Series of articles.) *Pflüg. Arch. f. d. ges. Physiol.*, 1914-18, **158-172**.

28. Tolman, R. C. Statistical mechanics. *Chem. Cat.* (New York), 1927.

29. Tracy, H. C. The development of motility and behavior reactions in the toadfish (Opsanus tan). *J. Comp. Neur.*, 1926, **40**, 253-360.

30. Tschukitscheff, I. P. Über den Mechanismus der Hungerbewegungen des Mageus. (I) Einfluss des "Satten" und "Hunger" Blutes auf die periodische Tätigkeit des Mageus. *Pflüg. Arch. f. d. ges. Physiol.*, 1929, **223**, 251-264.

31. Wada, T. An experimental study of hunger in its relation to activity. *Arch. Psychol.*, 1922, **8**, No. 57, 1-65.

32. Wang, G. H. The relation between "spontaneous" activity and oestrous cycle in the white rat. *Comp. Psychol. Monog.*, 1923, **2**, No. 6.

33. Wolf, E. Die Aktivität der Japonischen Tauzmaus und ihre rhythmische Verteilung. *Zsch. f. vergl. Physiol.*, 1930, **11**, 321-344.

Laboratory of General Physiology
Harvard University
Cambridge, Massachusetts

"L'ACTIVITÉ SPONTANÉE" CHEZ LES SOURIS

(Résumé)

Quand on fait tous les efforts possibles pour éliminer l'excitation près d'une souris, il a lieu toujours des mouvements—les soi-disant "mouvements spontanés." Ces mouvements ne se montrent pas par hasard, mais avec une forme dèfinie, c'est-à-dire, dans des périodes rapides de mouvement répétées régulièrement suivies de périodes de tranquillité; avec une température constante pour des intervalles d'observation durant environ deux heures, les périodes d'activité et de tranquillité sont d'une longueur constante, ayant une erreur probable moyenne, comme pourcentage de la moyenne de 5% pour l'activité, de 10% pour la tranquillité (à 21° C.). Les souris sont toujours poikilothermiques à l'âge de deux jours; la fréquence des cycles d'activité s'accroît avec la température à la manière d'une réaction chimique. Quand la fréquence moyenne de l'occurrence des périodes d'activité est construite selon l'équation d'Arrhenius, la valeur de μ a été de 25.300 calories. La durée des intervalles de tranquillité et d'activité change aussi comme fonction de température; quand leurs réciproques sont construites selon l'équation d'Arrhenius, la valeur de $\mu = 13.200$ est obtenue pour l'activité, $\mu = 30.900$ pour la tranquillité. On a suggéré comme hypothèse pratique que ces valeurs différentes indiquent que deux processus au moins ont été responsables du contrôle de l'activité et de la tranquillité. On a supposé que l'activité résulte quand la concentration d'une substance *A* se maintient au-dessus d'un niveau de seuil. La tranquillité commence quand *A* se trouve au-dessous de ce niveau, la vélocité de la formation et du déplacement étant contrôlée par des changements chimiques.

Les souris âgées de 40 à 90 jours montrent la même occurrence en cycles de périodes d'activité que la montrent les souris âgées de deux jours. La variation d'un cycle supposé base se montre en multiples de 15 minutes pour l'activité, en étapes de 24 minutes pour les périodes de tranquillité.

Les mécanismes que contrôlent l'activité locomotrice "spontanée" des mammifères peut avoir comme "centre" la région gastro-intestinale, puisque les mouvements de l'estomac vide font des corrélations avec l'incidence de l'activité du corps. Les témoignages des articles publiés sur les mouvements spontanés montrent que les mouvements d'un animal intact font une corrélation avec les vigoureuses "contractions de faim" de son estomac vide et sont déterminés par les mêmes.

STIER

"SPONTANE TÄTIGKEIT" BEI MÄUSEN

(Referat)

Wenn man die Erregung in der Umgebung einer Maus möglichst zu eliminieren sucht, lassen sich immer noch Bewegungen feststellen, die sogenannten "spontanen Bewegungen." Diese Bewegungen ereignen sich nicht zufällig, sondern in bestimmter Form, nämlich in regelmässig wiederkehrenden Bewegungswellen, denen Ruheperioden folgen; bei konstanter Temperatur, während einer Beobachtungszeit von etwa zwei Stunden, sind die Tätigkeits- und Ruheperioden von konstanter Länge, mit einer durchschnittlichen Fehlerquelle von 5% für die Tätigkeit und 10% für die Ruhe (21 C). Die Mäuse sind noch poikilothermisch im Alter von zwei Tagen; die Häufigkeit der Tätigkeitszyklen nimmt mit der Temperatur in der Art

und Weise chemischer Reaktionen zu. Wenn man die durchschnittliche Häufigkeit der wiederkehrenden Tätigkeitsperioden nach der Arrhenius Gleichung entwirft, so ergibt sich der Wert von μ als 25,300 Kalorien.

Die Dauer der Tätigkeits- und Ruheperioden ändert sich gleichfalls als Funktion der Temperatur; wenn man ihre Reziproken nach der Arrheniusgleichung entwirft, so erhält man 13,200 Kalorien als Wert μ für Tätigkeit, und 30,000 für Ruhe. Man schlägt als Arbeitshypothesis vor: diese verschiedenen Werte zeigen an, dass wenigstens zwei Prozesse für die Regulation der Tätigkeit und der Ruhe verantwortlich sind. Man nimmt an, dass Tätigkeit hervorgerufen wird, wenn die Konzentration einer Substanz *A* über einer bestimmten Schwelle steht. Ruhe beginnt, wenn *A* unter diese Schwelle erniedrigt wird, wobei die Geschwindigkeit der Bildung und des Verschwindens durch chemische Veränderungen reguliert werden. Vierzig bis neunzig Tage alte Mäuse zeigen dieselbe zyklische Wiederkehr der Tätigkeitsperioden wie die zwei Tage alten. Variationen eines angenommenen Grundzyklus erscheinen in Vielfachen von 15 Minuten für Tätigkeit, in Abständen von 24 Minuten fur Ruheperioden.

Der die spontane lokomotorische Tätigkeit von Säugetieren regulierende Mechanismus mag sein “Zentrum” im Verdauungssystem haben, da die Bewegungen des leeren Magens Bezeihungen zur Korpertätigkeit haben. Beweisstücke aus veröffentlichten Darstellungen bezüglich spontanen Bewegungen zeigen, dass die Bewegungen intakter Tiere Beziehungen haben zu und bestimmt sind durch starke Hungerkontraktionen des leeren Magens.

Stier

THE PROGRESSIVE INCREASE IN THE GEOTROPIC RESPONSE OF THE ANT *APHAENOGASTER**

From the Laboratory of General Physiology, Harvard University

T. Cunliffe Barnes and B. F. Skinner

Although various tropisms [cf. (1)] of animals are now well known (with the possible exception of chemotropism) there are several categories of movements which still appear to be rather anomalous. Spontaneous and intermittent activity, tropism reversals, "random" movements and variations in the degree of tropistic response require extended investigation. In the present paper certain somewhat novel features in the geotropic response of the ant *Aphaenogaster* are described, and partially explained in terms quite compatible with the traditional "forced movement" philosophy. One of us[1] has recently reported that typical geotropic orientation may be secured with ants on an inclined plane. In the course of those enquiries unusual changes in the geotropic sensitivity of *Aphaenogaster fulva* were discovered.

All our work with *Aphaenogaster* was carried on in a dark room under dim red illumination, for these ants are greatly affected by light. In bright light we were able to obtain definite phototropic paths which are of interest in connection with the studies of Loeb and of Santschi on ants. The former (3, p. 109) stated that he never found true heliotropism in the workers and the latter (4, p. 384) claimed that the disturbance produced in the behavior of ants by reflecting the sun's rays upon them with a mirror was not phototropism.

We have used the same simple technique which was found suitable with *Crematogaster* and *Lasius*. The insect is lifted gently on a small square of paper to the center of an inclined plane and allowed to crawl away in any direction. The path is followed with a soft pencil upon a sheet of paper on which the ant may secure a firm foot-hold at any inclination. As with *Crematogaster* and *Lasius*,

---

*Received for publication by W. J. Crozier of the Editorial Board, June 18, 1929.

[1] Barnes, T. C. *J. Gen. Psychol.*,

only the worker ants were used. Individual workers were studied over a wide range of angles of inclination (α) of the substratum, and others were tested at intervals for several days with a constant tilt of the plane. In both cases the Θ angle of geotropic orientation, measured as deviations from a line parallel to the horizontal axis of the plane exhibited a progressive increase. The method of measuring the Θ angle is given in Figure 3.

The rise in the gravity response may be seen in Figure 1 where the order in which the various records were taken is indicated on the curve. In this case the increase in the Θ angle is opposed by using lower and lower inclinations. At each point the angle of orientation represents the average of 40 measurements on 20 trails each about 15 centimeters in length. The heightened geotropic response, especially obvious between points 2 and 4, was finally overcome by the markedly decreased gravitational component at 5. The time at

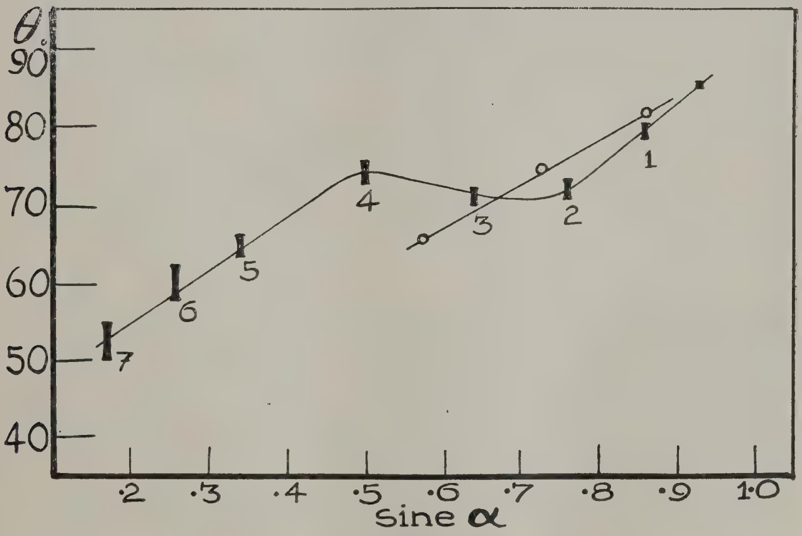


FIGURE 1

The intersection of the curve representing the geotropic orientation of *Aphaenogaster* at progressively lower angles and the line giving the response when the angle of inclination is progressively increased from low values. The numbers indicate the order in which the records were taken (*vide* Table 1). The bars are proportional to the probable errors. The line joining the dots gives the result of the ascending values of α for a second ant (*vide* Table 2).

TABLE 1

The Geotropic Response of *Aphaenogaster* No. 1 Tested at Various Times and Inclinations of the Plane

$n = 40$. Positive Geotropism

Trial No.	α (degrees)	θ *P.E.* (degrees)	Time
1	60°	79.4°±1.12	Sept. 14, 4-4:25 P.M.
2	50°	72.0°±1.39	4:36-4:52
3	40°	71.0°±1.24	5:06-5:18
4	30°	74.3°±1.53	5:27-5:56
5	20°	64.7°±1.45	10:23-10:29
6	15°	60.0°±2.15	Sept. 15, 11:10-11:31 A.M.
7	10°	52.7°±2.40	11:57-12:02
8	70°	85.4°±0.56	Sept. 17, 10:51-11:03 A.M.

TABLE 2

The Geotropic Response of *Aphaenogaster* No. 2 at Various Times and Inclinations (α) of the Plane

$n = 40$. Negative Geotropism

Trial No.	α (degrees)	θ *P.E.* (degrees)	Time
1	35	65.6±2.39	Sept. 18, 5:36 P.M. Sept. 19, 3.40 P.M,
2	45	74.4±1.41	Sept. 19, 4-4:30 P.M.
3	60	81.4±0.87	Sept. 24, 11:35 A.M. 4:42 P.M.

which each of these records was taken is given in Table 1. When the procedure is reversed and the first angles taken are the lower ones, a curve is obtained which intersects the first (the line indicated by circles in Figure 1). In Table 2 is recorded the first test made at $\alpha = 35°$, in which the θ angle is nearly 10 degrees lower than the (fourth) record in Table 1 at $\alpha = 30°$.

The effect is perhaps most striking when records are obtained with an *Aphaenogaster* on a surface of constant inclination over a period of several days. The individual is taken from the colony and kept in a watch glass with moist blotting paper without food. Such an ant dies in four to eight days if not returned to the nest. Figure 2 presents the progressive rise in the θ angle of nearly 10 degrees at 50° inclination of the substratum. It seemed quite possible that this rise might be interpreted as a species of geotropic "learning," and the following experiment was devised to throw light on the question. For a short time on each of three days an individual was permitted to crawl on a level surface for approximately the same distance as the usual inclined plane record and on the fourth day the geotropic orientation was tested on a 50° incline. For the following three days trails were secured again on the level and a final geotropic response was recorded at the 50° incline. The measurements represented graphically in Figure 4 lead one to suggest that the "training" under gravitational stimulation is not the chief factor involved in the progressive rise in Θ, for after three days with brief exercise on a horizontal surface the value has arisen to 68.5°,

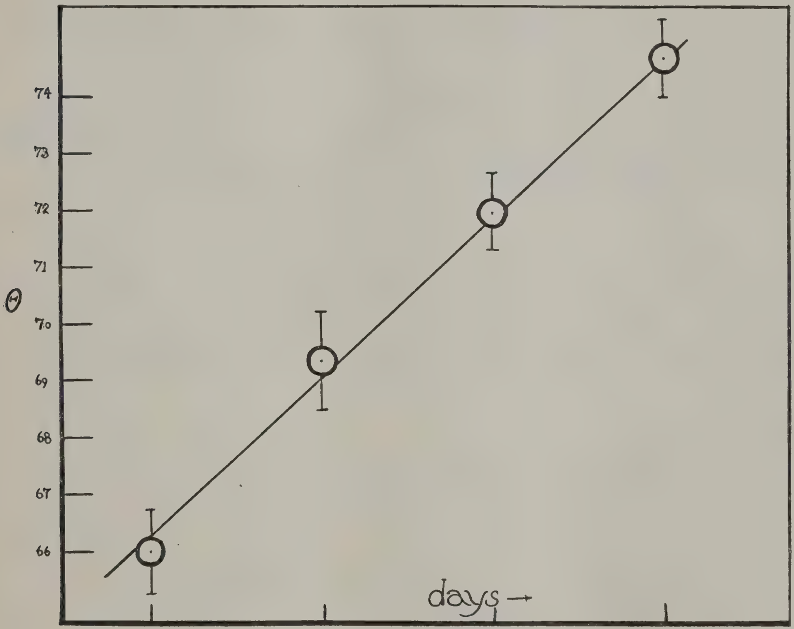


FIGURE 2

The progressive increase in the geotropic response of *Aphaenogaster fulva.* The orientation- or Θ-angle measured as deviations from a line at right angles to the longitudinal axis of the plane is tested once a day on a constant inclination of 65° on an inclined plane (the α angle). Each point represents the average of 40 determinations. The short vertical lines through the points are twice the probable errors. This ant eventually died, as do most isolated individuals.

and three more days of the same treatment result in an increase to 72.8°. This individual appeared to be stronger than the one whose record is given in Figure 2.

The crucial experiment suggested by Professor Crozier was carried out with a fresh individual. This ant was kept in isolation for three days without any noticeable creeping in the watch glass. On the fourth day the geotropic angle on a 50° inclination was found to be 80.4°! The possible effect of previous crawling on the incline or on the level was quite definitely excluded, but this ant exhibited the greatest increase in geotropic response. It is obviously important to keep in mind that these isolated individuals die if not returned to the

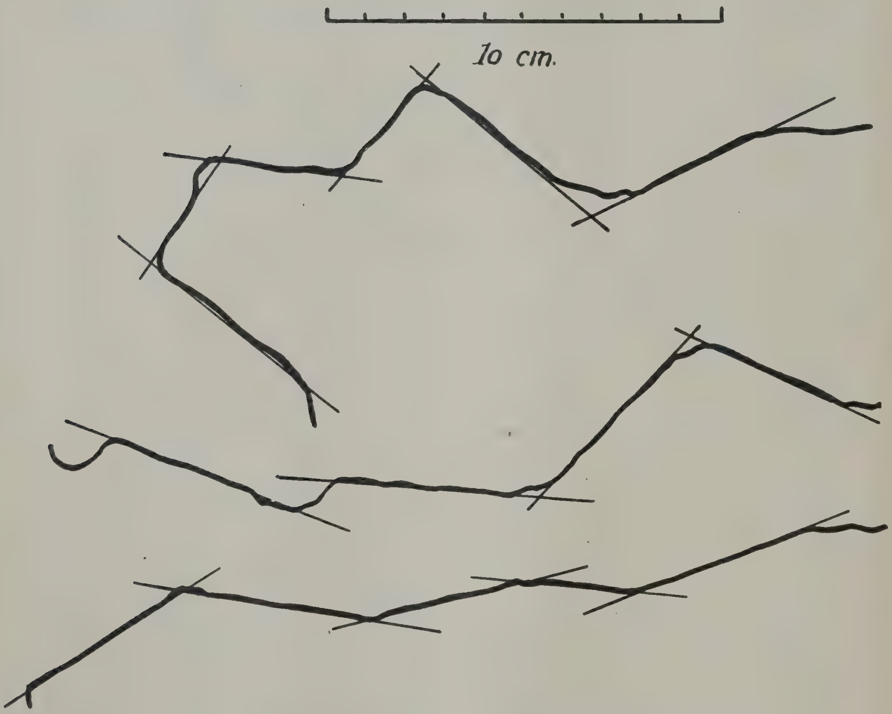


FIGURE 3

The method of measuring Θ angles (of the short straight lines superimposed on the recorded path of the ant).

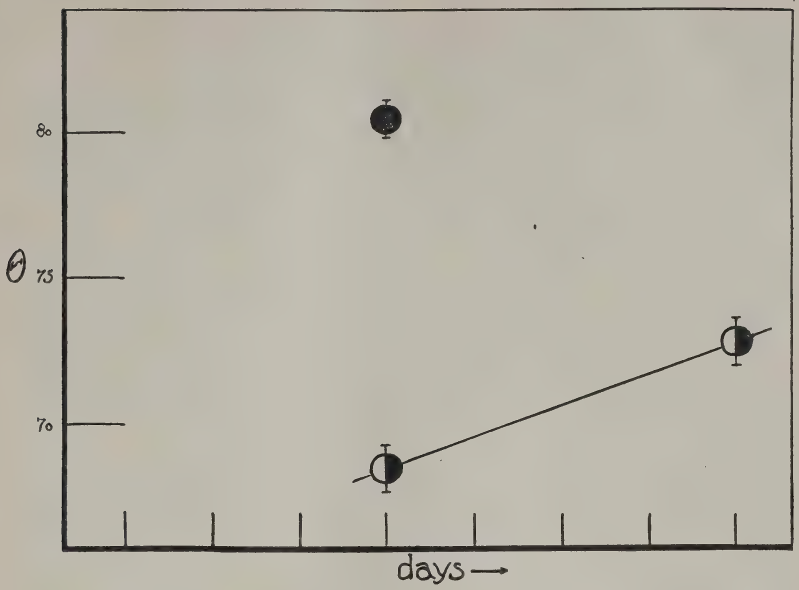


FIGURE 4

The rise of the angle of geotropic orientation with two individuals of *Aphaenogaster fulva.* The black and white points represent the rise in Θ from 68.52°±0.79° to 72.8°±0.80° in an *Aphaenogaster* exercised on a level surface at intervals and tested on a 65° inclination. The black point is the high Θ exhibited by an ant kept in isolation without crawling and then tested on a 65° slope. This value 80.45°±63° was the highest obtained at this slope.

colony, and, after some days confinement, certain individuals tend to drop from the inclined surface at high angles. One might cautiously postulate "muscular weakness" or some irreversible change proceeding to death, but in any event our evidence proves that we are dealing with some simple physiological change. Nor must it be forgotten that a change in posture[2] involving a displacement of the center of gravity would influence the extent of geotropic response, for we have previously established the "muscle tension theory" of geotropism in ants.

That the asymmetry in proprioceptive stimulation produced by the

---

[2] *Aphaenogaster fulva* has unusually long legs and one frequently observes maintained changes in the orientation and degree of flexion in the legs.

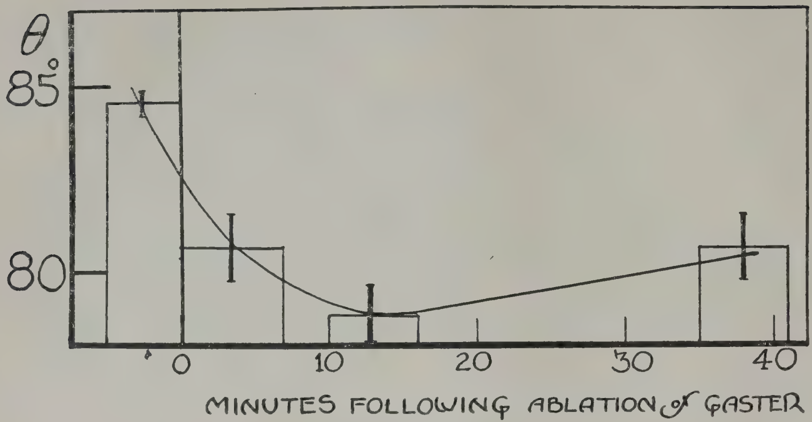


FIGURE 5

The decrease in the angle of geotropic response following ablation of the abdomen in *Aphaenogaster fulva.* $n = 40$; $a = 65°$.

weight of the ant on the anti-gravity muscles determines the geotropic response may be shown by ablation of the gaster. As indicated in Figure 5, the θ angle drops immediately. After a few preliminary trials the θ angle of this ant was measured at $a = 65°$. For each θ, 40 angles were averaged from about 10 paths each 18 centimeters long. The abdomen was then cut off and the θ value for the next series of paths became 80.6° as contrasted with 84.5° for the intact insect. In Table 3 the next four mean values are recorded, and it will be seen that the "progressive rise" in geotropic response described in the first part of this paper manifests itself again and raises the θ angle back to 84° on the second day. With *Aphaenogaster* No. 4, however, the low θ value was still exhibited 27 hours after the ablation of the gaster. With individual No. 1, whose abdomen was removed after it had given the extensive paths recorded in Table 1, a decrease of six degrees was obtained (Table 3).

Of course no ablation experiment is quite exempt from criticism, and the lower θ exhibited after the operation might follow from nervous disturbances. Although we regard this possibility as rather remote here, another method is available to demonstrate the muscle tension theory. Weights may be attached to the animal exhibiting

TABLE 3

THE EFFECT OF ABLATION OF THE GASTER ON THE ANGLE OF GEOTROPIC ORIENTATION IN *Aphaenogaster*

$n = 40$. The intact insect was used in Trials 1 and 9.

Trial No.	α (degrees)	Θ *P.E.* (degrees)	Time
Aphaenogaster No. 3 (negative geotropism)			
1	65	84.5±0.39	Nov. 15, 3:00 P.M.
2	65	80.6±0.92	3:05-3:12
3	65	78.8±0.85	3:15-3:21
4	65	80.6±0.92	3:43
5	65	81.6±0.71	Nov. 17, 10:50-11:00 A.M.
6	65	84.1±0.61	11:05-11:14 A.M.
Aphaenogaster No. 4 (positive and negative)			
7	65	81.5±0.90	Nov. 23, 2:48-2:55 P.M. (27 hours after ablation)
8	65	80.9±0.70	3:02-3:10 P.M.
Aphaenogaster No. 1 (positive geotropism)			
9	30	75.0±1.13	Sept. 17, 11:16-11:19 A.M.
10	30	68.7±1.48	11:22-11:42

geotropic orientation, and the effect on the Θ angle observed.[3] *Aphaenogaster* No. 2 exhibited a Θ angle of 69.1°±2.52° on an inclination of 35°. A pupa was placed before the ant on the plane. The insect picked up the pupa in its mandibles and continued its negatively geotropic orientation but at a Θ angle of 78.4±1.96°. The rise of nine degrees was produced by the weight of the pupa which, added to the worker's weight, elicited a greater proprioceptive stimulation in the supporting musculature of the latter. Each Θ represents an average of 15 measurements. The rise in Θ seems unusually great, and it would appear that the maintained act of gripping the pupa increased the tonus in the legs as well as in the mandibles.

The most effective way of causing a drop in the Θ angle, even with individuals which exhibit the progressive rise in Θ during a series of trials, is by ablation of the antennae. The path described on an inclined surface by an *Aphaenogaster* without antennae is quite sinuous and often rather erratic. The ant whose path after ablation of the antennae is recorded in Figure 6*A* gave a Θ angle of 77.7°±0.92° on an inclination of 65° to the horizontal when tested before the operation. The insect stopped crawling immediately after the antennae were cut off (Nov. 15, 1928, at 7 P.M.), but when placed on the plane ($\alpha = 65°$) two days later (Nov. 17, 11:30 A.M.) extensive crawling was exhibited. The θ angle had fallen from $77.7 \pm .92°$ to $61.1 \pm 2.09°$ ($n = 40$). The amplitude of the waves in the path increased, for we found a $\frac{1}{2}\lambda$ value of 24.4 ± 1.21 mm. before ablation and $19.0 \pm .65$ mm. after the loss of the antennae ($n = 26$). The number of waves per unit length of trail is also affected,

---

[3] Crozier and Pincus (2) showed that the angle of orientation in young rats increases about as the log of weight added at the root of the tail. In the geotropism work with rats it was not possible to test the Θ angle over a period of days as with the worker ants.

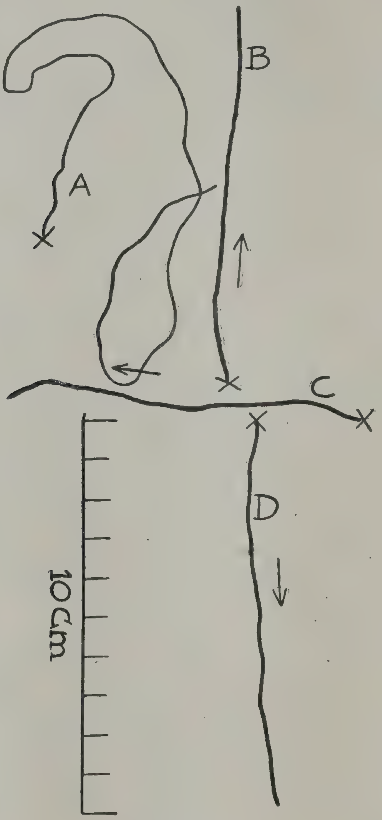


FIGURE 6

Paths described by geotropic *Aphaenogaster fulva;* all trails begin at "X."

A. Path of an ant without antennae. $\alpha = 65°$.

B, C, and *D.* The reversal of geotropism from negative to positive in an intact *Aphaenogaster.* $\alpha = 65°$.

which may be observed by measuring the distance between periodic turns or slight alternate changes in direction to right or left. Examination of 130 cm. of path in each condition showed that the intact ant moved 27.6 mm. between turns, which became 21.3 mm. (average) after ablation of the antennae. All these measurements were made with the same ant at a constant inclination of 65° to the horizontal.

The preservation of geotropic sensitivity after the loss of the antennae furnishes further support for the muscle tension theory of geotropic orientation, but our results indicate that the antennae-less insect is at least partially released from the gravitational θ-response. Perhaps the side-to-side tapping action of the antennae tends to check wide turning, as the $\frac{1}{2}\lambda$ value suggests. The ant executes balancing movements with the feelers, but it is difficult to attribute much importance to such actions from the standpoint of geotropism. It will be recalled that Wolf (5) found that antennae-less hive bees were not as greatly disturbed by rotation during transit as the intact bee. We hope to return to this unsettled point in future geotropism studies.

A reversal of geotropism occurred after the loss of the feelers in *Aphaenogaster*, from positive to negative, but such a change occurs at intervals in all geotropism records with this insect. In Figure 6, *B* represents one of 7 negative trails from an individual at $\alpha = 60°$; the next 8 trails were directly to the side, *C;* and the following 4 were quite positive, *D*. This occurred in less than 10 minutes with practically continuous crawling. The sign of the response is given for all paths in the tables.

SUMMARY

Workers of the ant *Aphaenogaster fulva* exhibit a progressive increase in the extent of their geotropic response on an inclined surface. This and several other modifications of the θ angle of geotropic orientation are shown to be generated by physiological causes which may be experimentally controlled.

REFERENCES

1. CROZIER, W. J. Tropisms. *J. Gen. Psychol.*, 1928, **1**, 213-238.
2. CROZIER, W. J., & PINCUS, G. The geotropic conduct of young rats. *J. Gen. Physiol.*, 1926, **10**, 257-269.
3. LOEB, J. Comparative physiology of the brain and comparative psychology. New York: Putnam, 1903. Pp. 309.

4. Santschi, F. Comment s'orient les fourmis. *Rev. suisse de zoöl.*, 1913, **21**, 347-426.
5. Wolf, E. Über das Heimkehrmögen der Bienen. *Zsch. f. vergl. Physiol.*, 1927, **6**, 221-254.

Laboratory of General Psychology
Harvard University
Cambridge, Massachusetts

LA CROISSANCE PROGRESSIVE DANS LA RÉPONSE GÉOTROPIQUE DE LA FOURMI APHAENOGASTER

(Résumé)

Les ouvreuses parmi les fourmis *Aphaenogaster fulva* montrent une croissance progressive dans l'étendue de leur réponse géotropique sur une surface plane. Cela et plusieurs autres modifications de θ, l'angle de l'orientation géotropique, se montrent produites par des causes physiologiques qu'on peut contrôler expérimentalement.

Barnes et Skinner

DIE PROGRESSIVE ZUNAHME DER GEOTROPISCHEN REAKTION BEI APHAENOGASTERAMEISEN

(Referat)

Es zeigt sich, dass bei Arbeiterameisen der Aphaenogaster fulva der Umfang ihrer geotropischen Reaktion auf einer schiefen Oberfläche progressiv zunimmt. Es ergibt sich, dass diese und andere Modifikationen von θ, dem Winkel der geotropischen Orientation durch physiologische Ursachen hervorgerufen werden, die experimentell geprüft werden können.

Barnes und Skinner

HYPNOSIS AND THE CONDITIONED REFLEX*[1]

From the Psychological Laboratory of the University of Wisconsin

Harold D. Scott

The existence of the phenomenon of post-hypnotic amnesia is a well-established fact. Strickler (8) recently found that when the subject was tested in the waking state by the *relearning* method following the trance learning of nonsense syllable equivalents of characters the amnesia amounted to only about 50%. Coors (2), using the *relearning* method in the waking state, found an amnesia of 45.56% for the stylus maze. No work has appeared, up to the present time, using a more basic behavior mechanism such as the conditioned reflex as the criterion of amnesia.

Pavlov (5) found that certain of his dogs upon being subjected to the conditioning process passed into a state in which no conditioned reflexes could be evoked; even unconditioned stimuli unless quite strong failed to give their normal response. He states that this condition greatly interfered with his experimentation. He also found an instance where a dog failed to respond to strong stimuli but did show responses which were larger than usual to weak stimuli (p. 271). Pavlov infers that the above conditions are hypnotic phenomena and have their analogy in man. At the same time he admits that it is quite difficult to draw a parallel between animals and human beings in the more complicated states of the trance (pp. 404 ff.) His use of catalepsy as a criterion of the presence of the hypnotic state in both animals and humans does not find very much support among the students of human hypnosis (2, 4).

In regard to the state of the nervous system at the time of conditioning, Pavlov believes that it must be "alert" in order to establish the reflex with any degree of readiness (6, p. 28). If we regard

---

*Recommended by Clark L. Hull and received for publication by Carl Murchison of the Editorial Board, February 19, 1930.

[1] This is the seventh of a series of studies in the psychology of suggestion by research workers under the direction of Professor Clark L. Hull. The present study is a preliminary report of a group of studies by the author on the subject of hypnosis. He wishes to express his indebtedness to Professor Hull for his invaluable suggestions and criticisms during the progress of the experiment.

the trance as a kind of sleep, as Pavlov does, this would seem to argue against the formation of the conditioned reflex in the trance state except possibly in a long drawn-out procedure with a very slight degree of success.

The observations of Pavlov suggest the following problems in relation to human subjects: (*a*) Will conditioned reflexes set up in the waking state function in a succeeding trance state? (*b*) Is it possible to set up conditioned reflexes while the subject is in a state of hypnosis? (*c*) If the second problem is found to yield an affirmative solution, will the conditioned reflex function after the subject is brought out of the trance, i.e., is the conditioned reflex subject to post-hypnotic amnesia the same as ordinary memories? (*d*) Providing either Problems *a* and *b* yield positive results, what will be the relation of the strength of the conditioned reaction when tested in the following state as compared with its strength in the state in which it was set up? (*e*) Judging from the data obtained in these experiments, what is the relation of human hypnosis to what Pavlov calls hypnosis in dogs?

Should it be found that reflexes established in the waking state pass over into the trance state quite consistently, and, should it be possible to set them up while the subject was in a somnambulistic trance (one for which there is a post-hypnotic amnesia), these findings would severely shake Pavlov's contentions regarding the nature of hynosis and its relation to the conditioning process.

The primary aim of the present experiment was to investigate the phenomenon of post-hypnotic amnesia by attempting to set up a conditioned defense finger reflex to a buzzer, with an electric shock as the unconditioned stimulus, while the subject was in the hypnotic trance. If this proved successful, the next step was to test for the presence of the conditioned reaction in the waking state fifteen minutes after establishment. A simultaneous record was taken of the breathing, heart action, and the galvanic skin reflex, in addition to the finger response. The records of the heart action and the galvanic skin reflex are not reported because of the inadequacy of the records secured. Another group of subjects were to be conditioned in the waking state to the same stimuli as controls in order to determine the relative susceptibility of trance and waking subjects to the conditioning tendency.

The shock stimulator used was modeled on the plan of the one used by Watson (10, Figure 2, p. 105). It consisted of a metal plate, upon which the heel of the right hand rested, forming one electrode; a small metal saddle, shaped to receive the middle finger, formed the other. Another saddle carved from wood rested on the top of the finger. When the finger moved in response to a stimulus this movable saddle pressed upward against the surface of a tambour placed above it, which in turn transferred the impulse pneumatically through rubber tubing to the recording tambour at the kymograph. The shock strength was controlled by means of a 50-watt lamp and an adjustable rheostat set in series with the regular 110-volt alternating house current.

The stimulus keys were of the ordinary telegraph variety. These keys were arranged so that magnetic signal markers recorded on the smoked paper drum the points at which the stimuli were introduced. An ordinary buzzer gave the conditioned stimulus. The breathing record was obtained by means of a Sumner pneumograph. A Zimmerman pendulum time clock set to make contacts at five-second intervals, and a connected magnetic time-marker, gave a parallel time line.

The subject in the trance experiment was prepared in an adjoining room. Introspections were taken on his general condition; he was blindfolded, and taken into the experimental room. Here no instructions of any kind were given. He was seated in a chair beside the shock stimulator which was mounted on a table. The second finger of the right hand was inserted between the finger electrode and the recording saddle, and the tambour moved down by means of a rack and pinion device to make proper contact. The pneumograph was next adjusted about the subject, after ascertaining whether he breathed thoracically or abdominally, and the air pressure equalized. The galvanic and sphygmomanometer attachments were made on the fingers of the left hand and the left arm respectively in the cases where they were used. A sample section of one of the smoked paper records is reproduced in Plate 1.

The kymograph was started and allowed to run for about two minutes while the subject became adapted to the apparatus, the experimenter taking his position at the stimulus controls behind the subject. From three to five isolated buzzer stimuli were given, about three-fourths of a second in length, at intervals of approximately one minute between stimuli. This was done to obtain a normal control on the reactions in the waking post-trance period. The subject was

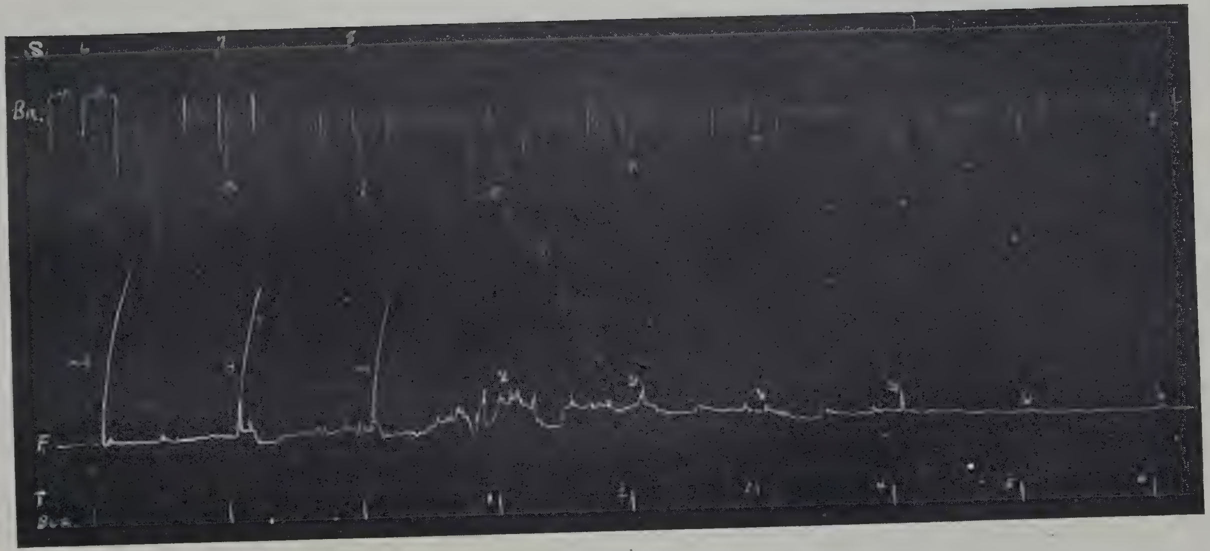


PLATE 1

This figure shows a sample section of one of the smoked paper records of a reflex established in the trance. The last three reinforcements, numbered 6, 7, and 8 at the upper left, are shown for comparison with the conditioned responses which follow. This subject was somewhat nervous during the latter part of the experiment which makes the finger-tracing somewhat jagged. Only eight reinforcements were needed to establish the conditioned reaction. The arrows point to the significant responses during the conditioning and the test periods.

Legend: *S*—the electric shock. *Br.*—the breathing reactions. *A*—(above the breathing tracing) marks off a fifteen-second period preceding the stimulus. *B*—(ibid) marks off a fifteen-second period following the stimulus. *F*—the finger-defense reaction. *T*—time line (in five-second intervals). *Buz.*—the buzzer stimulus.

then returned to the adjoining room where he wrote out what he had noticed about himself during the experiment.

He was then hypnotized, blindfolded, returned to the experimental room, and readjusted into the apparatus. The interval between the preceding phase and this one was 15 minutes. From three to six isolated buzzer stimuli were again given as a trance control, in the same manner as was done previously in the waking state.

The buzzer was now reinforced with the shock which followed it each time after an interval of about one-fifth second. These combined stimuli were given with approximately one minute between pairs, with the exception of two subjects where the interval was varied from 45 to 75 seconds to see whether the constant time interval was a major factor in the conditioning.

The subject was ordinarily tested for the conditioning after at least 10 buzzer-shock combinations had been given. If he showed a strong finger reaction to the combined stimuli he was tested earlier. Where the conditioned finger reflex showed up quite strongly for three successive tests it was considered established. If it did not show an appreciable reaction the establishment was continued. The subject was then tested at intervals later in the experiment. In all cases the reflex was further reinforced after its first showing in order to strengthen it. After the conditioned response showed satisfactorily the subject was removed from the apparatus to the adjoining room. Those subjects with whom conversations could be carried on in the trance were questioned as to their experiences and their answers were recorded. The blindfold was now removed and the subject allowed to rest for a few minutes.

The subject was next wakened and the blindfold placed on him. He was returned to the experimental room 15 minutes after the termination of the trance session and tested for the presence of the conditioned response to the isolated buzzer. On this occasion at least three isolated buzzer stimulations were given each subject.

Records were obtained of the reactions of ten subjects for the hypnotic, and nine for the waking establishment. The subjects were all male undergraduates, members of the author's classes in introductory psychology. Of the ten trance subjects, measurements were taken on eight for the finger reaction—two of the records being misplaced before the final computations could be made. However, a previous inspection had shown that both these subjects gave conditioned finger and breathing responses. The pneumograph developed a leak on one of the remaining subjects, leaving seven on which breathing measurements were obtained.

RESULTS[2]

The finger-defense reaction was scored by measuring in fiftieth inches the height, or depth, of the tracing on the smoked paper record from the beginning of a movement to its peak. This reaction showed much variation in its behavior both as to the kind and strength of the response. In some cases the finger straightened out upon the contact instead of rising, thus causing a drop in the tracing. Where the individual jerked the hand completely out of the apparatus, as sometimes occurred, an arbitrary value of —25 was given the reaction. These negative values were treated as positive in computing the composite indices for the finger. The composite reactions representing the eight subjects measured are shown in Figure 1. In compiling the tables from which this graph, as well as the others, was derived, the stimuli were arranged in separate categories (waking buzzer pre-test, trance buzzer pre-test, buzzer-shock establishment, trance buzzer tests for evidence of conditioning, and waking buzzer post-tests for conditioning), and numbered from one to the highest numbered stimulus given in that group. The measurements, in the case of the finger-defense, and the ratios of the breathing ratios were entered by subjects in their appropriate tables. The mean value for each reaction on each stimulus was found and the results plotted in the composite graphs.

Of the ten subjects used in the trance conditioning, all except two showed consistent conditioned finger-withdrawal (defense) reactions in that state. One of these two jerked his hand out of the shock apparatus on the 13th buzzer-shock stimulation—the arm becoming quite rigid. The experimenter commanded relaxation which followed immediately. After a pause of ten minutes the subject was tested in the trance state and gave a single finger reaction out of three buzzer tests. The other subject developed a conditioned muscle twitch in the upper right arm so that his record of the conditioning tests while in the trance shows a continuous series of zeros which accounts for the apparent extinction tendencies on the 11th and 12th isolated buzzer test stimulations shown in Figure 1. This subject was the only one given 12 test stimulations and only one

---

[2] For the detailed measurements of each subject see (7).

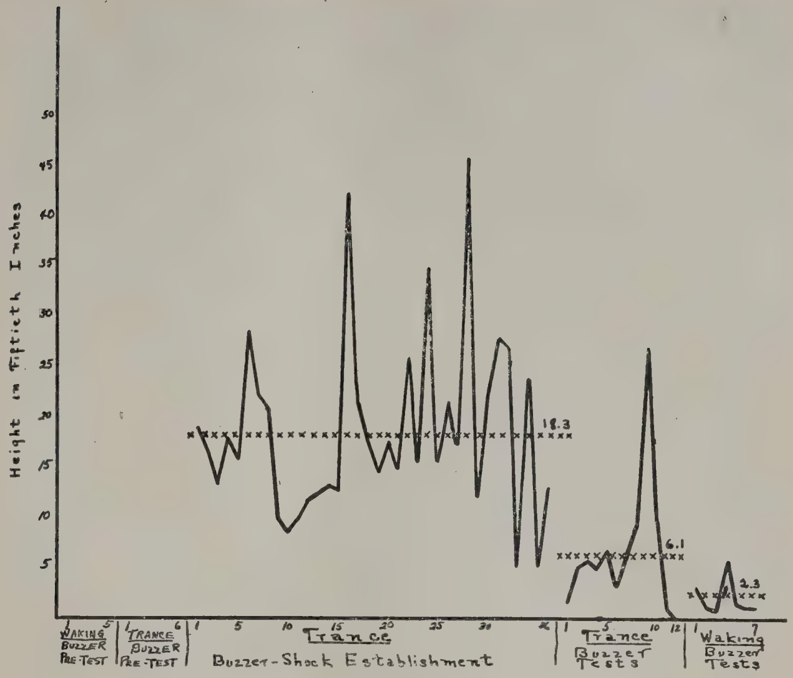


FIGURE 1

Graph Showing the Composite Finger-Defense Reactions of Eight Trance Subjects

xxxxx represents the mean height of the reactions for the period of the experiment

An interval of approximately 15 minutes occurred between the last trance conditioning test and the first waking buzzer post-test

other was given 11, so the last two values are not exactly representative. The high peak in the trance conditioning tests on the ninth buzzer test is due to the tendency of one subject to show a violent increase in his conditioned reaction during the tests given instead of an extinction.[3]

---

[3] Because of the small number of subjects and the more or less exploratory nature of the investigation the graphs presented contain fairly marked sampling errors. However they are presented to show general tendencies found. Since the subjects were not all given the same number of test or establishment stimuli the graph values represent increasingly smaller numbers of subjects near the end of each division of a record.

Four different indices were based on the breathing reactions. The amplitudes of the inspiration heights in the 15-second period preceding a stimulus, termed the *A* phase, were measured in fiftieths of an inch by taking the perpendicular distance from the beginning of each inspiration to the maximum height attained by it. These measurements were likewise taken of the 15-second period following the stimulus, the *B* phase. The total effect of both inspirational and expirational oscillations during each phase was measured with

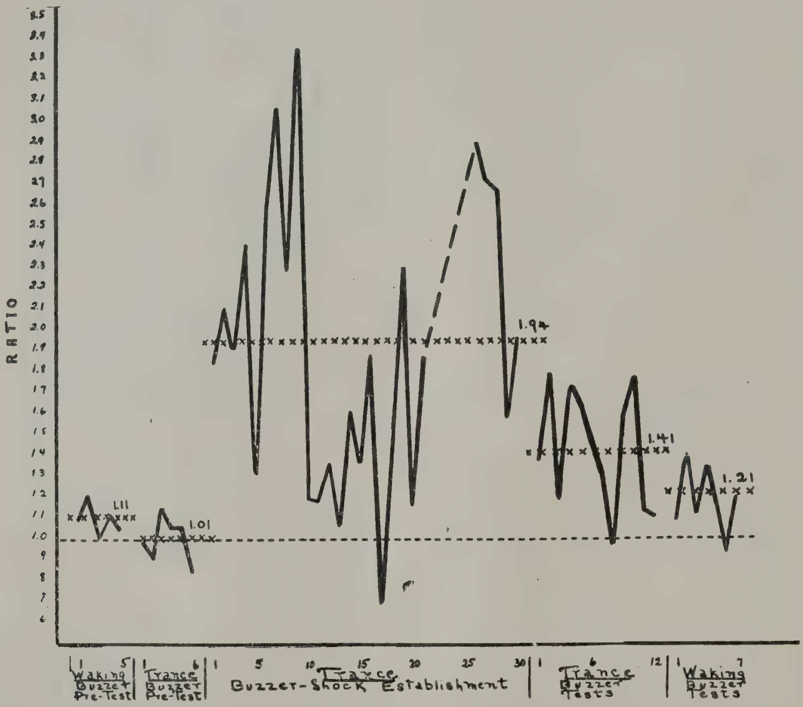


FIGURE 2

GRAPH SHOWING THE COMPOSITE $\frac{\text{MEAN AMPLITUDE } B}{\text{MEAN AMPLITUDE } A}$ RATIOS ON THE BREATHING OF SEVEN TRANCE SUBJECTS

xxxxx represents the mean value of the ratios for each period of the experiment

----- represents the point of zero disturbance in the breathing

No measurements were made on the 22nd, 23rd, 24th, and 25th combined buzzer-shock stimuli for those subjects carried beyond that point in the conditioning process.

the Hull linear oscillometer (3). These measurements were in terms of degrees of the circular recording wheels, each degree being equivalent to 1/40 inch.

The first index of breathing conditioning which we shall consider was derived from the main heights of the *A* and *B* phases by finding the ratio of the phase following to that of the one preceding the stimulus. If it was consistently greater than 1.00 this would indicate an increased amplitude in the *B* phase, or if less than 1.00,

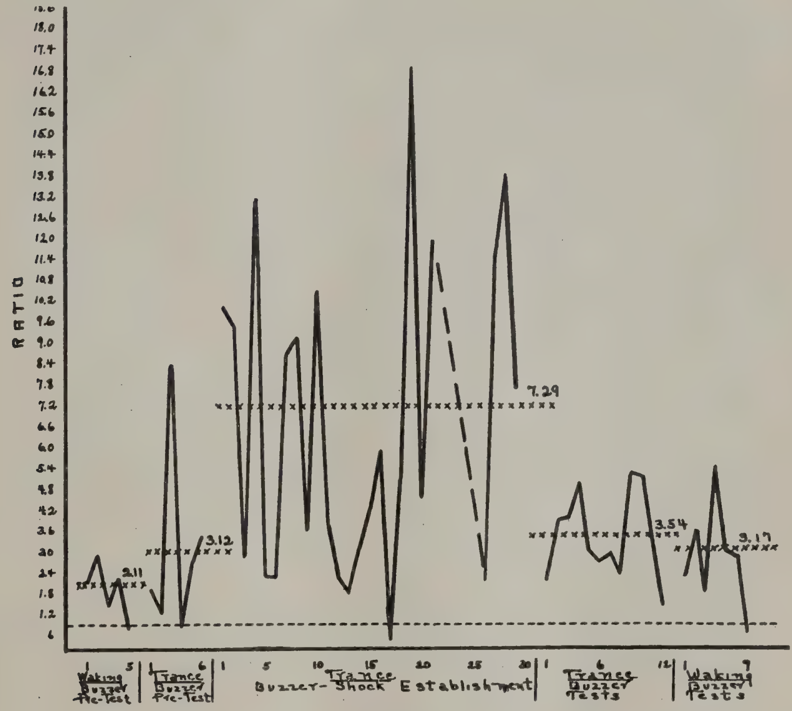


FIGURE 3

Graph Showing the Composite $\frac{\text{Mean Deviation } B}{\text{Mean Deviation } A}$ Ratios on the Breathing for Seven Trance Subjects

xxxxx represents the mean value of the ratios for each period of the experiment

----- represents the point of zero disturbance in the breathing

No measurements were made on the 22nd, 23rd, 24th, and 25th combined buzzer-shock stimuli for those subjects carried beyond that point in the conditioning process.

a decreased amplitude due to the stimulus. These ratios appear in Figure 2.

To determine the effect of the stimulus upon the variability of the inspiration heights the ratio of the mean deviation of the *B* phase to that of the *A* phase for each stimulation was computed. Figure 3 shows the behavior of this index. A value consistently greater than 1.00 would indicate an increased disturbance in rhythm due to

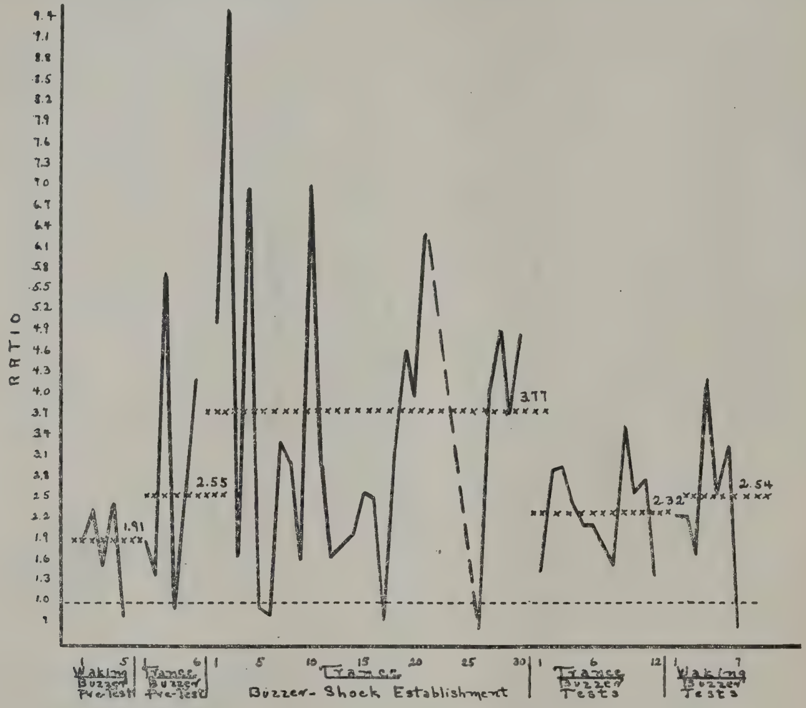


FIGURE 4

GRAPH SHOWING THE COMPOSITE $\frac{\text{Mean Deviation Percentage Mean } B}{\text{Mean Deviation Percentage Mean } A}$ RATIOS ON THE BREATHING FOR SEVEN TRANCE SUBJECTS

xxxxx represents the mean value of the ratios for each period of the experiment

– – – – – represents the point of zero disturbance in the breathing

No measurements were made on the 22nd, 23rd, 24th, and 25th combined buzzer-shock stimuli for those subjects carried beyond that point in the conditioning process.

the stimulus; a value of less than 1.00 would show the opposite effect.

A third method was to express the mean deviation of each phase in percentages of the mean from which it was derived. A ratio was obtained between these derived figures as in the above cases. This was done to try to obtain an expression of the conditioning in situations where the mean deviation, though quite small, was large when compared with the mean amplitude value, or the reverse of these conditions in some cases.

Lastly the ratio between the oscillations of the *B* to the *A* phase was taken for each stimulus. The oscillation measurements were made by tracing the breathing curves with the small pointer on the oscillometer. The movements of the pointer caused the recording wheels to revolve. When the tracing of the phase was completed the total of the inspiration and expiration movements were added giving a total oscillation measurement on that phase.

The breathing was measured on all the isolated buzzer stimuli given a subject. With the combined buzzer-shocks, measurements were taken for the three reinforcements just preceding each isolated buzzer test group, with the exception of one or two instances where more than three were measured. The buzzer-shock measurements were for the purpose of obtaining a comparison with the response to the isolted buzzer when given following the reinforcement.

The mean values, in the pre-tests, of all the breathing indices on the composite graphs (Figures 2, 3, 4, and 5), show that the buzzer already possessed a slight excitatory potency. This is what one would expect from the normal emotional value which a sudden noise possesses for most individuals. In the waking pre-tests for all indices, and in the trance pre-tests for the *mean height* and *oscillation* ratios, the values show a tapering off toward the neutral point—what Pavlov would call a "neutralizing effect." The distortion of the trance pre-test values of the *mean deviation* and *mean deviation percentage of the mean* ratios (Figures 3 and 4) is due to a sharp sigh occurring in the *B* phase of stimuli Nos. 3 and 6 of one subject. Otherwise the same general tendency would probably have appeared.

The mean values of the ratios of the buzzer-shock series rise sharply above the values for the preceding isolated buzzer stimuli. This is due to the greater excitatory effect of the shocks, which is exactly what one would expect. There is a prolonged effect from the shock which could often be detected in the *A* phase of the next stimulus combination. The typical reaction was a deep inspiration followed by erratic breathing if the stimulus was given while an inspiration was in progress; or if given on an expiratory movement it showed a short, quick inspiration followed by a period of deeper, erratic, breathing. The trance buzzer tests for conditioning showed this same type of reaction. This latter tendency often accounted for a smaller mean amplitude in the *B* phase than occurred in the *A* phase just preceding it. The mean index values for the conditioning tests were much higher than those preceding the reinforcements.

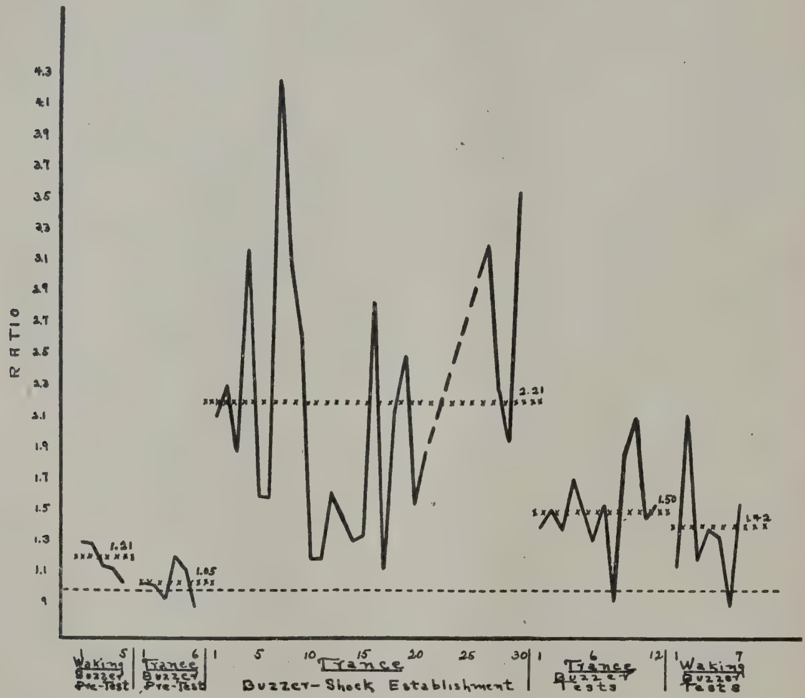


FIGURE 5

Graph Showing the Composite $\frac{\text{Oscillation } B}{\text{Oscillation } A}$ Ratios on the Breathing for Seven Trance Subjects

xxxxx represents the mean value of the ratios for each period of the experiment

––––– represents the point of zero disturbance in the breathing

No measurements were taken on the 22nd, 23rd, 24th, and 25th combined buzzer-shock stimuli for those subjects carried beyond that point in the conditioning process.

What appears to be a practice effect in the conditioned response appears in the trance conditioning tests in Figures 1, 3, 4, and 5. The second isolated buzzer test for conditioning in both trance and waking conditioning tests shows a stronger reaction than the first on the indices for the finger-defense and breathing with the exception of the waking post-test for the finger and the *mean deviation percentage of mean* ratio in the same series. This tendency for the reaction to increase in value for the first two or three tests with the isolated conditioned stimulus before experimental extinction began to show was also noticed by Switzer (9) in his study of the conditioned lid reflex. It supports the view that previous to the onset of experimental extinction the conditioned reflex shows an increase through use.

All of the nine trance subjects who were tested for the conditioned reflex in the waking state following the trance showed both a conditioned finger-defense and conditioned breathing response. These conditioned reactions were smaller in size, for the most part, than were the ones during the trance period just preceding (see Figures 1, 2, 3, 4, and 5). The mean indices of all the ratios of the breathing were consistently higher in the conditioning post-tests than they were in the waking pre-tests. The indices all show the effects of the conditioning passing from the trance into the following waking state although none of the subjects was aware that he had been conditioned. Each subject remarked about his reactions which occurred in the waking post-tests, expressing surprise at the changes in behavior from that of the waking pre-tests. None became aware of the nature of the process until told about it at the completion of the experiments. It thus appears that, if the principle of post-hypnotic amnesia operates on the conditioned reflex at all, its influence must be distinctly limited.

It would seem from this work and that of Strickler and Coors, cited at the beginning of this article, that post-hypnotic amnesia does not signify an inability to reinstate activity performed during the trance but rather an inability to tell about it. This would seem to have far-reaching significance in the field of mental hygiene.

A comparison of the records from the trance and waking (those conditioned while in the waking state) subjects shows the former much more susceptible to conditioning of the finger-defense reaction than the latter. All ten of the trance subjects showed a conditioned finger-response—exception should probably not be made for the individual who showed the shoulder twitch in the trance but who gave the conditioned finger reaction in the waking state following. These comparisons are summarized in Table 1.

TABLE 1

Showing the Relative Ease of Establishment of the Finger-Defense Reaction in Trance and Waking States

	No. tested	No. showing conditioning	Percentage of no. tested showing cond. resp.	Av. no. comb. buzzer-shocks necessary to produce consistent reaction	Subjects not showing cond. finger reaction No.	Av. no. of comb. B.-S. stim. given
Trance estab.	10	10	100	14.2	0	
Waking estab.	9	5	55.6	26.6	4	30.3

Note: The average number of shocks on the trance established group was computed on the eight records available for final computation.

This evidence clearly contradicts Pavlov's contentions that the conditioned reflexes do not function in the hypnotic state—providing the ease of conditioning in the trance proves to be due primarily to the presence of the trance state. Until suitable controls are run by conditioning persons who are susceptible to the trance while in the waking state and then comparing the results with the behavior of establishments on trance-resistant subjects and with subjects conditioned while in the trance, the contradiction must remain somewhat hypothetical. However, since two of the subjects not showing a conditioning, among those subjected to the reinforcement in the waking state, were excellent trance subjects, it would appear that the conditioned reflex can be established more readily in the trance state. This latter observation casts some doubt upon the hypothesis that subjects susceptible to the trance are also more susceptible to conditioning.

An interesting and somewhat illuminating group of observations was made during the experimentation which seems to show a kind of redintegrative effect. Subject L. M., on a waking retest for the conditioned reaction four days after the original trance conditioning, jerked his hand back as it was being inserted into the electrodes and exclaimed, "I had a shock!" He did this again on the trance test which followed. On being questioned in the trance he reported having received a shock. The writer tested the electrodes after the first report, and both before and after the latter one and found no current present. This subject also reported a feeling of extreme coldness and numbness occurring at the sound of the buzzer in his report following the waking tests. One other subject noticed this feeling of shock when being adjusted into the apparatus for a waking retest, after having been conditioned in the trance state. Neither of these two had any waking recollection that they had received any shocks during the establishment period, though of course they really had. One subject, conditioned in the waking state, also reported this feeling upon being adjusted for a retest of the conditioning, but said that it was only momentary.

A more striking observation was reported by two subjects, one conditioned in the trance, and the other in the waking state. The trance subject's report is quoted in detail.

During Trance

Experimenter's question (after conditioned reaction had appeared) "What happened since you have been sitting here?"

Subject's answer (in a faltering manner) "I got a shock when I heard the buzzer. There was a pain in my arm and a burn on the end of my finger. I caught my breath and felt upset all over."

(*Q*) "Did you notice anything when the buzzer sounded by itself?"

(*A*) "It felt the same way as when the shock came. I felt pain too."

In the waking post-test which followed he made the following report: "At the sound of the buzzer I was conscious of a jerk in my right hand. In addition I caught my breath. I have come to fear the sound of the buzzer and do not like it. I expect something worse than really comes from the sound. I seem to feel conscious of some pain in my finger each time."

The other individual, conditioned in the waking state, said that he felt a distinct tingling in the finger which rested upon the electrode at the sound of the buzzer upon the first conditioning test after shock reinforcements had been given. As additional reinforcements were given the feeling became more intense on the occasion of the isolated buzzer stimuli. He claimed, however, that the feeling was not as intense as that of the actual shock, but this is not to be wondered at since conditioned responses under experimental conditions rarely equal the strength of the response upon which they are built. Seven individuals reported distinct fear responses, shivers, visceral disturbances, coldness, etc., to the buzzer after it had been associated with the shock.

These observations give us some insight into another phase of conditioning, namely, the subject's feeling the unconditioned stimulus even though it is no longer impinging on him. It would suggest that psychologically the unconditioned stimulus does not drop out, as we are wont to have it do in our symbolic representations of the conditioning process, but remains potent when the conditioned stimulus is applied alone. It appears to be the more *real* the greater the strength of the conditioned response. This suggests that the essential basis of sensory experience is *action;* and that, if an organism should react in all ways as it is accustomed to react to a given stimulus, it will *feel* that stimulus, regardless of what the *real* stimulus was, i. e., what other organisms stimulated by it at the same time would report it to be.

Summary and Conclusions

1) Conditioned reactions are set up in the trance with greater readiness than in the waking state. This is contrary to the implications of Pavlov's theory of hypnosis. The indication is that the state which Pavlov calls hypnosis in his dogs is not the same as that called hypnosis in humans.

2) The trance-conditioned subjects showed a consistent conditioned finger-defense response after an average of 14.2 combined buzzer-shock stimuli as against a showing of five out of nine waking subjects, the latter showing the conditioning after an average of 26.6 combinations. The remaining four waking subjects failed to show the conditioned reaction after an average of 30.3 reinforcements. In other words, the trance state appears to facilitate conditioning, instead of interfering with it as Pavlov seems to imply.

3) Conditioned reflexes set up in the trance may be evoked in the succeeding waking state in spite of the fact that the subjects can recall nothing of the trance events, including the process of conditioning. The results show that the conditioned reflex is, at least to a considerable extent, free from the waking amnesic tendency ordinarily associated with trance events. It would tend to show that what is commonly termed post-hypnotic amnesia is not a total amnesia but is a temporary inhibition of the habit systems connected with motor speech which prevents the subject from being able to

make a symbolic report of the activity although he does reinact the activity learned in the trance when the stimulus conditions are repeated.

4) The conditioned reactions are much smaller in magnitude, or intensity, than the unconditioned reactions. This is in harmony with previous work on the conditioned reflex.

5) An inspection of successive tests for the existence of the conditioning tendency in the trance shows a relatively small reaction on the first stimulus followed by increasing values for the next two or three reactions. After this brief practice effect, the magnitude of the responses gradually declines, according to the principle of experimental extinction.

6) The appearance of a *feeling* of the unconditioned stimulus when it is not actually present suggests that the basis of sensory experience is *action*. Psychologically the unconditioned, or primary stimulus, appears to continue to function even though objectively it is not present.

References

1. Bramwell, J. M. Hypnotism, its history, practice and theory. (3rd ed.) London: Rider, 1921. Pp. xvi+480.
2. Coors, D. A determination of the density of post-hypnotic amnesia for the stylus maze. A. B. Thesis, University of Wisconsin, 1928. On file in the Library of the University of Wisconsin.
3. Hull, C. L. An instrument for summating the oscillations of a line. *J. Exper. Psychol.,* 1929, **12**, 359-361.
4. Moll, A. Hypnotism. (Trans. by A. F. Hopkirk.) (3rd ed.) New York: Scribner's, 1913.
5. Pavlov, I. P. Conditioned reflexes. (Trans. by G. V. Anrep.) London: Oxford Univ. Press, 1927. Pp. xv+430.
6. ————. Lectures on conditioned reflexes. (Trans. by W. Horsley Gantt.) New York: International, 1928. Pp. 391.
7. Scott, H. D. The conditioned reflex and hypnosis. M. A. Thesis, University of Wisconsin, 1929. On file in the Library of the University of Wisconsin.
8. Strickler, C. B. A quantitative study of post-hypnotic amnesia. *J. Abn. & Soc. Psychol.,* 1929, **24**, 108-119.
9. Switzer, S. A. A backward conditioning of the lid reflex. *J. Exper. Psychol.,* 1930, **13**, 76-97.
10. Watson, J. B. The place of the conditioned-reflex in psychology. *Psychol. Rev.,* 1916, **23**, 89-116.

University of Wisconsin
Madison, Wisconsin

L'HYPNOSE ET LE RÉFLEXE CONDITIONNEL

(Résumé)

On a fait cette expérience dans le but d'étudier le phénomène de l'amnésie post-hypnotique en essayant d'établir un réflexe conditionnel de défense du doigt à une sonnette électrique (avec un choc électrique comme le stimulus non conditionnel) pendant que le sujet a été dans la catalepsie hypnotique et en employant des tests pour trouver la réaction conditionnelle dans l'état de réveil quinze minutes après l'établissement.

Les résultats ont montré que les réactions conditionnelles sont établies dans l'état de catalepsie plus facilement que dans l'état de réveil (ce qui est contraire aux implications de la théorie de l'hypnose de Pavlov). De telles réactions ont été évoquées dans l'état suivant de réveil malgré le fait que les sujets n'ont pu rien se rappeler des événements de la catalepsie. En accord avec les résultats d'autres études, les réactions conditionnelles se sont montrées de moins de grandeur et d'intensité que les réactions non conditionnelles. Le principe de l'extinction expérimentale aussi a été opératif. L'apparence d'un *sentiment* d'un stimulus non conditionnel quand celui-ci a été absent suggère que la base de l'expérience sensitoire soit *l'action.* Psychologiquement le stimulus non conditionnel ou primaire semble entrer en fonction même bien qu'il ne soit pas objectivement présent.

SCOTT

HYPNOSE UND BEDINGTER REFLEX

(Referat)

Das Hauptziel dieses Experimentes war die Untersuchung der nachhypnotischen Amnesie, und zwar durch die Erstellung eines bedingten Fingerreflexes auf den Reiz eines Summers (mit elektrischem Schlag als unbedingtem Reiz) im hypnotischen Transzustand und durch die Prüfung der bedingten Reaktion fünfzehn Minuten nach der Erstellung im wachen Zustand.

Die Ergebnisse zeigen, dass bedingte Reaktionen im Transzustand leichter hergestellt werden als im wachen Zustand (eine Tatsache, die mit Folgerungen der Pavlovschen Theorie über Hypnose im Widerspruch steht). Solche Reaktionen wurden in nach einander folgenden Wachezuständen hervorgerufen, trotzdem sich die Versuchspersonen der Begebenheiten des Transzustandes nicht erinnern konnten. Man fand in Übereinstimmung mit den Ergebnissen anderer Studien, dass die bedingten Reaktionen weniger gross und intensiv sind als die unbedingten. Das Prinzip der experimentalen Auslöschung hat sich ebenfalls bestätigt. Die Tatsache, dass das Gefühl des unbedingten Reizes auftritt, auch wenn er nicht vorhanden war, weist darauf hin, dass die Unterlage der sensorischen Erfahrung wirksam ist. Psychologisch scheint also der unbedingte oder primäre Reiz wirksam zu sein, wenn er objektiv nicht vorhanden ist.

SCOTT

EMOTIONAL REACTIONS TO STRONG STIMULI*

From the Psychological Laboratory of the University of Wisconsin

J. E. CASTER[1]

INTRODUCTION

Attempts to discriminate between the emotions by analysis of physiological changes have been most discouraging. Studies of changes in respiration, circulation, and electrical resistance thus far have failed to reveal characteristic physical patterns. At present we rely upon introspections and reports of internal observations of our subjects for the differentiation of their states. The dependence of this technique on the discriminative powers of the subject, and the emphasis upon the subjective rather than the objective method seriously weakens its validity. It is conceivable that further fruitful experimentation in this field must await the researches of physiological chemistry. However, a more immediately promising method is to submit the experimental results to more exact quantitative and mathematical treatment.

The rhythmic nature of all functional activities of the autonomic nervous system suggests the possibility of studying the emotions from that angle. If emotions are indeed chaotic disturbances of visceral tonicity, as has been supposed by some, they should reveal something of this in the increased irregularity of the behavior rhythms. We have at our command various statistical methods for determining variability which may be applied extensively to the records of emotional disturbances. This may reveal new quantitative indices which will be of value in the study of individual reactions to various kinds of strong emotional stimuli.

Blatz (2) studied some of the changes in fear aroused in his subjects by the sudden and unexpected precipitation backward of the chair in which they were seated. His investigation deals chiefly with cardiac and electrical disturbances, the latter recorded by a string-galvanometer. Since this apparatus registers changes in the electro-

---

*Recommended by Clark L. Hull and received for publication by Carl Murchison of the Editorial Board, October 29, 1929.

[1] Grateful acknowledgment is hereby made to Professor Clark L. Hull for many suggestions relative to the problem.

motive force of the heart muscle, Blatz assumes that "the degree of deflection is a proportional measure of the intensity of the change."

In the interpretation of his records, rate, rhythm, and force of the cardiac changes were examined. The regularity of rhythm was determined by comparing the duration of the successive heart cycles measured in twentieths of a second. The amount of deviation from the average was then taken as the measure of the degree of irregularity.

Respiration was recorded by an electric pneumograph device, the tracings of which indicated the beginning and the duration of inspiration and expiration only. Blatz measured the regularity of respiration by computing the difference in rate of the consecutive groups of tracings as well as the inspiration-expiration ratio.

APPARATUS

In setting up our own experiment an attempt was made to maintain in our subjects as casual an attitude as possible.

We were seeking to observe a change in the regularity of the rhythmic pattern of certain organic activities. To establish the existence of such a change the degree of the regularity of the rhythm following a stimulation must be compared with that of a state which is approximately normal. By the latter we mean the activity of the subject when at rest. Naturally we must reduce apprehension to the minimum. Perfection in this regard is perhaps impossible under the conditions imposed by laboratory methods, but it may be approached.[2]

Twenty young men students from the several colleges of the University, members of quiz sections in the course in elementary psychology, were selected to act as subjects. Most of them were extremely naïve as regards the apparatus employed, and all were in complete ignorance as to the nature of the experiment.

The research was conducted in a small darkroom of the psychology laboratories. The walls are painted a flat black. Ventilation is maintained by access to an air shaft, foul air being removed by a mild suction. The temperature was kept constant by thermostatic control. Upon a low shelf extending the entire length of one side of this room rested the various pieces of apparatus. During the course of the experiments the room was lighted by an incandescent lamp enclosed within two conical lamp shades secured together in such fashion that a small beam of light was directed upon the writing styluses. The subject sat with his back to the instruments about four feet from *E*. His chair was placed in such a position as to enable him to rest his arm on the shelf. The two electrodes used were of the "Ostwald" type. For the construction of this electrode a large test tube was fused at its base with a smaller tube of the same length, the whole being U-shaped. A small length of platinum wire was enclosed in the solid fused portion so that the ends protruded, one into the larger tube and one into the smaller. The two electrodes were embedded, three inches apart, in a small sand box which was secured to the arm of a chair. Within the large test tube a physiological saline solution, mercurous-chloride paste and mercury made contact by means of the platinum wire with a mercury column in the smaller tubes. From the latter, wires of copper, amalgamated and properly insulated, led off to the galvanometer. The circuit was closed by the insertion of the index and ring finger of the subject into the large tubes.[3]

---

[2] The inability to remove apprehension without lying to the subject is indicated by the fact that, following the experiment, all of the subjects confessed to a degree of expectancy. In spite of this fact, several subjects actually dozed in the chair during the rest period.

To insure that a constant area of skin should be exposed to the solution a length of adhesive tape was wound about the fingers which were then immersed somewhat past the edge of the tape. Before immersion a bit of cotton saturated in ether was rubbed over the fingers to remove secretions of the sebaceous glands. The psychogalvanometer was of the Stoelting-Wechsler cabinet type. The beam of reflected light from the galvanometer suspension as it moved back and forth over the translucent scale of the instrument was traced by a device especially designed for this purpose, and illustrated in Figure 1.

Respiratory movements of the thoracic cage were recorded by a Sumner pneumograph connected by tubing to a Marey tambour. This tambour was inverted so that the ascending stroke of each breath cycle represents inspiration. Variations in the rate and amplitude of the pulse were registered by a Meek-Erlanger recording sphygomanometer. A leather cuff enclosing a rubber bladder was secured by straps to the arm over the brachial artery. No attempt was made in this experiment to determine systolic and diastolic pressures. With practically all subjects a pressure of from 85 to 95 mm. Hg. was found to give a tracing of sufficient amplitude for our purpose. It was observed that subjects with large fleshy arms required a higher pressure to secure a satisfactory tracing.

---

[3] These electrodes are substantially the same as those designed by Professor F. N. Pierce of the Department of Physical Chemistry at the University of Iowa. We are indebted to Dr. C. A. Ruckmick of that university for a description of them.

Time was registered in five-second intervals by an electro-magnetic marker attached in series to a master time clock which hung in an adjoining room. A second electro-magnetic marker in circuit with a key was employed to indicate the instant of stimulation. A second key closed the circuit through an electrode which was secured to the right ankle of the subject. This electrode was formed by attaching two nickel discs to opposite sides of a leather wrist strap. These discs were moistened by a saline solution before attachment. An alternating current regulated as to intensity by a rheostat was used in producing the shock.

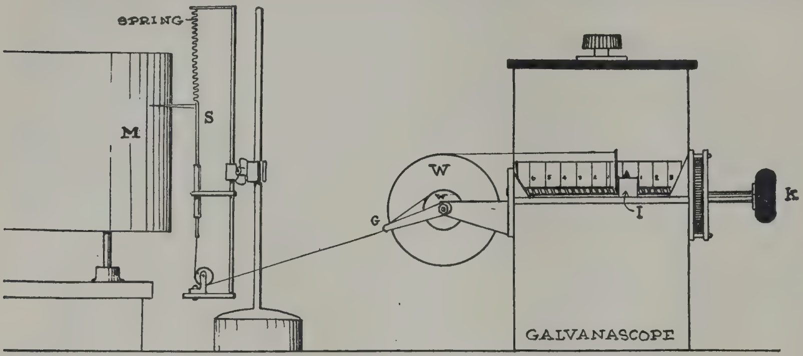


FIGURE 1

By the turning of knob (*K*) a spur gear imparts to the worm shaft a speed several times as fast. By this means the indicator (*I*) is made to follow the beam of light. From the indicator a thread passes to a large brass wheel (*W*), to which is soldered a smaller wheel (*w*) of one-third the circumference of *W*. (Each of these wheels has a grooved rim to retain the thread.) From the smaller wheel a thread passes through the guiding arm (*G*) and under a pulley connecting with the spring stylus (*S*). This stylus avoids the arc common to the lever-stylus usually employed. The tracing reduced one-third from the distance actually traversed by the indicator is then easily compared with the other changes employed. The stylus portion of the apparatus was designed by Dr. Clark L. Hull. (*M*) indicates the position of the idler drum of the extension kymograph.

The records were taken on smoked paper revolving about an extension kymograph. A sample of a portion of one record showing the relative arrangement of the styluses is reproduced in Plate 1. This records the response of Subject 11 to the shock stimulus.

PROCEDURE

On arrival at the laboratory the subject was escorted to the experimental room where his coat was removed and his sleeves rolled up above the elbows. He was then permitted to rest undisturbed in the chair for five minutes. *E* then instructed him as follows: "This experiment is designed to record certain of your normal reactions. There is nothing to fear. Relax and think of nothing but rest. During the course of the experiment you will close the eyes and avoid, as much as possible, voluntary movements of your arms."

E now instructed *S* to insert his fingers in the electrodes. The beam of light on the translucent scale of the galvanascope was brought to zero by the small adjustment screw of the galvanometer suspension. The arm cuff of the sphygmomanometer was placed on the arm and the pressure pumped up to the proper level. The stimulating electrode was strapped to the ankle. After the attachment of the pneumograph and the adjustment of the styluses, the kymograph was started.

After a control series of from 10 to 12 inspirations, *E* pressed the shocking key. Reactions to this stimulus were generally rather pronounced, including movements of the leg and upper trunk. These were permitted to subside, fluctuations of the galvanic skin reflex as indicated by the beam of light being constantly followed by manipulation of the knob by *E*. During the period of subsidence *E* reassured *S* that this stimulus would not be used again and urged him to relax. On the appearance of relative normality *E* pressed a pin firmly against *S*'s neck but in such a way as to avoid puncturing the skin. After the subsidence of disturbances from this source, a pistol was drawn from its place of concealment and fired directly back of *S*. The report was exceptionally loud. The kymograph was permitted to turn until the various tracings assumed an appearance of comparative stability.

After a removal of the apparatus from the subject's body, he was given a card upon which was printed the following list.

1) Very intense
2) Intense
3) Moderate
4) Slight
5) Untouched

He was instructed to recall his subjective reactions to all three stimuli and rate them on this five-point scale. It is significant that in nearly every case the report was given with little hesitation. The subject's subjective grading was then inserted at the proper point in the smoked-paper record.

RESULTS

1) *The Galvanic Skin Reflex.* In recording the galvanic skin reflex we were interested only in indications of comparative disturbance. Since we did not secure individual resistances, it was impossible to make any but the roughest comparisons between subjects [see (5)]. Nevertheless the tendencies in each record are apparent enough to have significance.

The method of recording the fluctuations of the galvanometer has been described.[4] The appearance of a typical tracing of the recording stylus for this index is shown in Plate 1. It will be noted that after a latent period of five seconds, Subject 11 suffered a sudden reduction in skin resistance arising from application of the shock. The disturbance is apparently slow in subsiding, for the line does not return to its former level within the limits of the record. The latent period which we find in this subject is also found with all members of the group, varying from one to five seconds between the instant of stimulation and the initial movement of the galvanometer mirror.

Records of the galvanic reaction seem to indicate the greater stimulative value of the shot. To both the shot and the shock stimuli all subjects showed a positive reaction as indicated by a sharp rise of the line traced on the record. Subjects showing a considerable disturbance from the shock and shot stimuli, as a rule, showed a noticeable but comparatively slight reaction when the pin was pressed against the skin. Few of the subjects knew the exact nature of this stimulus and several confessed after the experiment that they were anticipating a second shock, this time on the back of the neck. Fifteen per cent of the subjects gave no perceptible reaction to the pin stimulus. It is interesting that these individuals also gave but minimal galvanic reactions to the other two stimuli. Seventy-five per

---

[4] In measuring the fluctuations of resistance the relative displacement of the light on the reading scale of the instrument was used, i.e., the fluctuations of the light above its position just previous to the application of the stimulus.

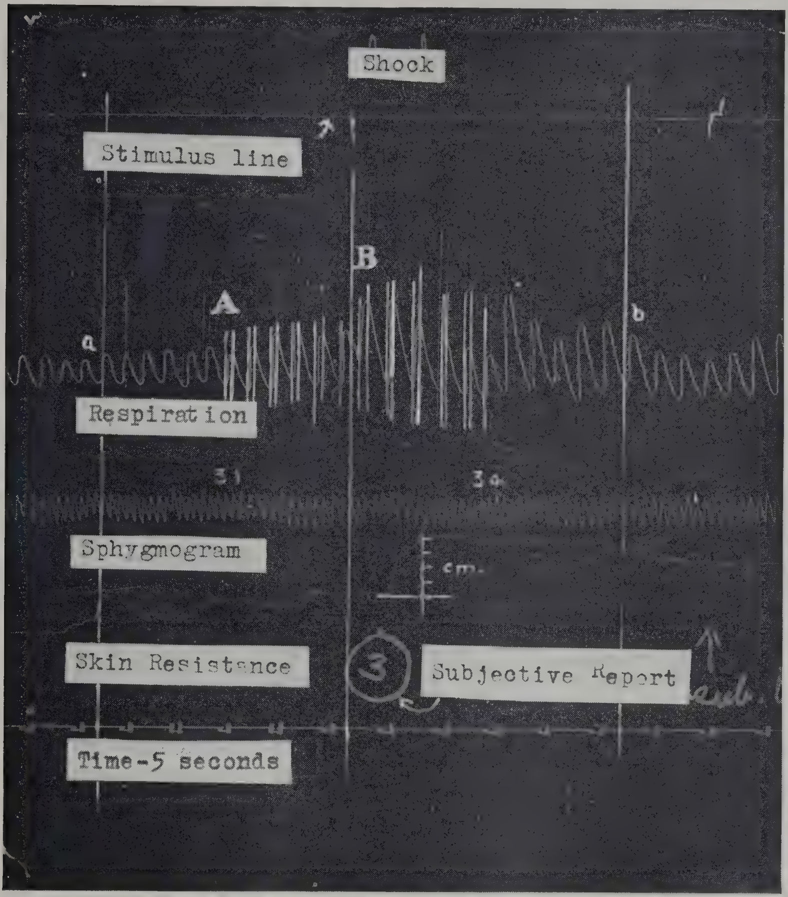


PLATE 1

A sample record showing the disturbances in Subject 11 as a result of the shock stimulus. By means of the time line, intervals of 25 seconds were marked off immediately preceding and immediately following the instant of stimulation. The distance from (*a*) to (*b*) therefore covers a period of 50 seconds. To secure measurements of the inspiration-expiration ratio, vertical lines are drawn through the highest and lowest points of each of the five respiratory cycles just preceding and following the point of stimulation. By measuring the horizontal distances between the lines, we secure data from which to compute the *I/E* ratio. This operation was facilitated by inverting the recording stylus so that each up-stroke on the record marks an inspiration.

cent of the subjects showed a gradual increase in skin resistance during the progress of the experiment. This was indicated on the records by the upward slope of the tracing from left to right. The galvanic tracing, as we have seen, rose sharply after each stimulus, then fell slowly, assuming a horizontal position somewhat above its original level.

While there still exists some doubt as to the exact nature of the psychogalvanic reflex, our experiments indicate that alterations in the resistance of the skin, as a rule, vary in proportion to the intensity of the stimulus as experienced by the subject. That these results agree with the findings from the other indices used will be shown.

2) *The Pulse.* It is well known that the heart is sensitive to reflex retardation and acceleration through its nerve supply. Heart rate has long been observed as a significant index of emotional disturbances. In this experiment we confined our attention to periods covering 25 seconds before and after each stimulus on the assumption that such disturbances would be sufficiently obvious in so brief an interval. Thirty-five per cent of the subjects showed an increase of heart rate as a result of the shock. In 10% of the cases no alteration was detected, while the remainder evidenced varying degrees of retardation of the pulse. After the firing of the shot 60% of the subjects responded with varying degrees of increase in rate. The pulse of one subject was retarded while the others showed no disturbance. Reactions to the pin stimulus were so diversified in regard to pulse rate that no typical response can be isolated. It is worthy of observation, however, that the majority of the subjects were rather consistent in regard to their heart rate. Those who suffered a minor alteration from the shock were but slightly disturbed by the other stimuli. This is consistent with the galvanic disturbances described above.

The amplitude of the pulse as recorded by our apparatus is altered by sensory stimulation. This variable lends itself easily to quantitative measurement when any suitable recording device is employed. In our experiment the rotation of the kymograph was at such a speed as to give a fairly compact but clear-cut sphygmograph. With the aid of a small transparent scale graduated in fiftieths-of-an-inch we could measure with accuracy the amplitude of each pulse wave.[5]

For the sake of clearness we refer again to Plate 1. With the aid of the time line a period of exactly 25 seconds before the stimulus was spaced off by erection of line (*a*) perpendicular to the time line. A similar line (*b*) was erected after a period of exactly 25 seconds from the stimulus which occurred at the point through which passes line (*b*). We now proceeded to measure the amplitude of each pulse wave in the two periods. As they were measured they were entered in columns as shown in Table 1. After completion of the measurements, the sums and means were computed. With these data we next determined the deviations from the mean amplitude of each separate pulse wave. The average of these columns yields an index which we have termed the *mean deviation of pulse amplitude.* Table 1 reveals the fact that Subject 11 showed an increase in this index of nearly 100% as the result of stimulation by the electric shock.

The percentage of subjects showing an increase in the *mean deviation of pulse amplitude,* together with the average for the index expressed in percentage of the pre-stimulus value for each of the three stimuli, is shown in Table 2.

It appears that a rough correlation exists between the increase of heart rate from strong stimuli and the variability of the amplitude of the various cycles revealed by our index of mean deviations. However, that these factors may vary independently was observed in several cases, e.g., with no increase in heart rate, a decided increase in the variability of amplitude was sometimes detected.

It will be recalled that the heart rate showed the greatest average increase as a result of the shot stimulus. To the shock 55% of the subjects responded with a slowing of rate, with 10% showing no change within the period examined. The ambiguity of response to the shock stimulus and also to the pin suggest that there are cases

---

[5] For investigations of deviations from the mean in the *rate* of the pulse we have been successful in employing an eyepiece micrometer with the low power of a compound microscope. This method is somewhat laborious but may reveal significant changes impossible of detection by the unaided eye. The scale of the micrometer glass, as seen through the eyepiece is adjusted in such a manner as to permit measurement of the horizontal distances between crests of the pulse waves. Means for any convenient period may be secured. Average deviations from the mean serve as a valuable index of variability. We employed this technique only occasionally in the present investigation, so that significant results cannot be reported.

TABLE 1

This table illustrates the method employed in tabulating the measurements of amplitude of the pulse beats and of deriving from them the index known as the *mean deviation of pulse amplitude.* These measurements were made directly from the pulse tracing of Subject 11, Plate I.

	Amplitude of the separate pulse cycles		Deviations from mean amplitude of the pulse cycles	
	Before stimulation	After stimulation	Before stimulation	After stimulation
	8.0	8.0	0.4	0.8
	8.0	10.0	0.4	1.2
	9.0	7.0	0.6	1.8
	7.5	6.5	0.9	2.3
	9.0	6.5	0.6	2.3
	8.0	7.0	0.4	1.8
	8.0	7.0	0.4	1.8
	9.0	9.0	0.6	0.2
	8.0	10.0	0.4	1.2
	8.0	9.0	0.4	0.2
	9.0	10.0	0.6	1.2
	8.0	8.0	0.4	0.8
	8.0	9.0	0.4	0.2
	9.0	10.0	0.6	1.2
	8.0	9.0	0.4	0.2
	8.0	7.0	0.4	1.8
	9.0	7.0	0.6	1.8
	8.0	9.0	0.4	0.2
	9.0	10.0	0.6	1.2
	8.0	9.0	0.4	0.2
	8.0	8.0	0.4	0.8
	8.0	9.0	0.4	0.2
	8.0	10.0	0.4	1.2
	8.0	9.0	0.4	0.2
	9.5	8.0	1.1	0.8
	8.0	10.0	0.4	1.2
	9.0	9.0	0.6	0.2
	10.0	10.0	1.6	1.2
	8.0	11.0	0.4	2.2
	9.0	9.0	0.6	0.2
	9.0	10.0	0.6	1.2
		10.0		1.2
		10.0		1.2
		8.0		0.8
Sums	261.0	299.0	16.8	35.0
Means	8.4	8.8	.54	1.03

TABLE 2

Showing the percentage of the 20 subjects who responded with an increase in the index, *mean deviation of pulse amplitude,* together with the percentage increase which the average index for the group as a whole shows for the period following the stimulus over the pre-stimulus value.

	Shock	Pin	Shot
Percentage of subjects showing increase in *mean deviation of pulse amplitude*	75	60	90
Response to stimulus expressed in increase over pre-stimulus value	41	21	100

where the heart rate may fail as a reliable indicator of emotional reaction. Contrasted with heart rate, we find that the *mean deviation of pulse amplitude* shows very definite alterations when expressed in percentages. These are especially marked with the percentage of increase in value over the pre-stimulus value. The shot appears here to be the strongest stimulus. This is significant in the light of the subjective estimations of the intensity value of the various stimuli which will be considered below.

The results of the present investigation indicate that pulse amplitude and, particularly, the *mean deviation of pulse amplitude* may be a very useful indicator of emotional reaction, fully as useful, if not more so, than heart rate, which is so generally employed.

3) *Respiration.* The extreme sensitivity of the respiratory center to sensory stimulation leads us to seek here for significant variations which may lend themselves to quantitative measurement. We might expect that strong sensory stimuli, such as were employed in the present investigation, would result not only in alterations in the amplitude of inspiration but also in an increased variability of amplitude. Can we not apply the statistical measure of the mean deviation to the breath cycle as we did with the pulse wave?

The measurement of any of the respiratory indices covering a considerable period after the stimulus would probably tend to mask the immediate disturbance resulting from it. We accordingly confined our attention to a comparison of the five cycles preceding with the five immediately following each of the stimuli.[6]

---

[6] In our records of respiration we encountered a difficulty which interferes somewhat with the exact measurement of the tracings, i.e., the arc of the stylus on deep inspiration becomes so great as to show serious lateral distortion. This could be avoided by employing a stylus the movements of which are strictly vertical. In spite of this handicap our measurements show rather consistent results.

In finding the amplitude of the respiratory cycle, we measured the vertical distance from the lowest point of the expiratory phase to the crest immediately following. It is noticeable that the inspiratory stroke was in most cases nearly vertical. Using once more our transparent scale which was employed in the measurement of the pulse waves, we measured the amplitude of each of the five cycles preceding and the five following the vertical line (*b*) (see Plate I). The results of these measurements were entered in Columns 1 and 2, respectively, of Table 3. Here it is seen that the mean amplitude after the stimulus, 20, is distinctly greater than before, 15.2. That this increase is a distinct tendency seems to be suggested by the fact that 75% of the subjects showed an increase resulting from the shock stimulus, 80% from the shot, and 35% from the pin. To obtain the average increase or decrease of this index over the pre-stimulus value is to reveal the general tendency. There was found for the shock 11.4% average increase in amplitude for 20 subjects, and for the shot 20% increase. The pin, which has already revealed itself as a weaker stimulus, showed a decrease of 1.6% over the pre-stimulus value. These results tend to show that strong sensory stimuli may provoke, for a brief interval at least, deeper inspirations.

There is a possibility that in addition to this increase in amplitude of respiration there exists an increased variability which may be used as an additional index, of value as a quantitative indicator of emotion. Accordingly the mean deviation of the depth of inspiration was found for the period before stimulation and separately for the period following the stimulus. For the former the difference is found between each entry in Column 1 and its mean, 15.2 (Table 3). These are entered in Column 3. The arithmetical average of these deviations (2.16) is known as the mean deviation. Column 4 shows the corresponding computations for the mean deviations following stimulation. We discovered a substantial increase in the variability, 2.16 to 3.2, or nearly 50%. For this index 75% of the subjects show an increase from the shock, 90% for the shot, and 40% for the pin. The percentage of increase of this index over the pre-stimulus value expressed in percentage is large for the two stronger stimuli, 69 for the shock and 225 for the shot. The pin reveals itself again as a weak stimulus, showing an increase of only .5% over the pre-stimulus value. The variability of amplitude is thus shown to be as valuable as amplitude itself in a study of disturbances of respiration.

An interesting question here presents itself. What is the average

TABLE 3

The measurements shown in this table have been taken from the breathing records reproduced in Plate 1. It is given in detail for the purpose of illustrating the exact method of deriving certain previously unused values from these original measurements.

Depth of breathing

	(1) Depth of inspiration before stimulus	(2) Depth of inspiration after stimulus	(3) Deviation of depth before stimulus	(4) Deviation of depth after stimulus
1	10	20	5.2	0
2	15	28	.2	8
3	18	19	2.8	1
4	16	18	.8	2
5	17	15	1.8	5
Sums	76	100	10.8	16
Means	15.2	20	2.16	3.2

Breaths preceding stimulus horizontal distances (time)

	(5) Inspiration	(6) Expiration	(7) Total inspiration-expiration cycle	(8) *I/E* ratio	(9) Deviation total cycle	(10) Deviation ratio
1	2.5	5.0	7.5	.50	.3	.09
2	2.5	5.5	8.0	.45	.2	.04
3	2.0	5.5	7.5	.36	.3	.05
4	2.0	5.5	7.5	.36	.3	.05
5	2.5	6.0	8.5	.40	.7	.01
Sums	11.5	27.5	39.0	2.07	1.8	.24
Means	2.5	5.5	7.8	.41	.36	.048

Breaths following stimulus horizontal distances (time)

	(11) Inspiration	(12) Expiration	(13) Total inspiration-expiration cycle	(14) *I/E* ratio	(15) Deviation total cycle	(16) Deviation ratio
1	2.5	6.5	9.0	.38	.7	.02
2	2.5	6.0	8.5	.41	.2	.05
3	3.0	7.5	10.5	.40	2.2	.04
4	1.0	6.0	7.0	.16	1.3	.20
5	2.0	4.5	6.5	.44	1.8	.08
Sums	11.0	30.5	41.5	1.79	6.2	.39
Means	2.2	6.1	8.3	.36	1.24	.078

amplitude of the consecutive respirations from the 20 subjects taken as a group? Stated in another way, we are interested in discovering any general tendency relative to breathing which might be revealed by such a computation. In Figure 2 we show the composite curves constructed from the average amplitude of each consecutive inspiration before and after the stimulus.

The characteristic alteration in amplitude of inspiration appears to be shown by the curves for the shock and the shot. The first inspiration after the stimulus is deep. This sudden aeration of the lungs is followed by a very shallow inspiration, succeeded again by a gradual increase in depth, with inspiration at the termination of five cycles after the stimulus deeper than normal. Slight disturbance of inspiration occurs from the pin stimulus.

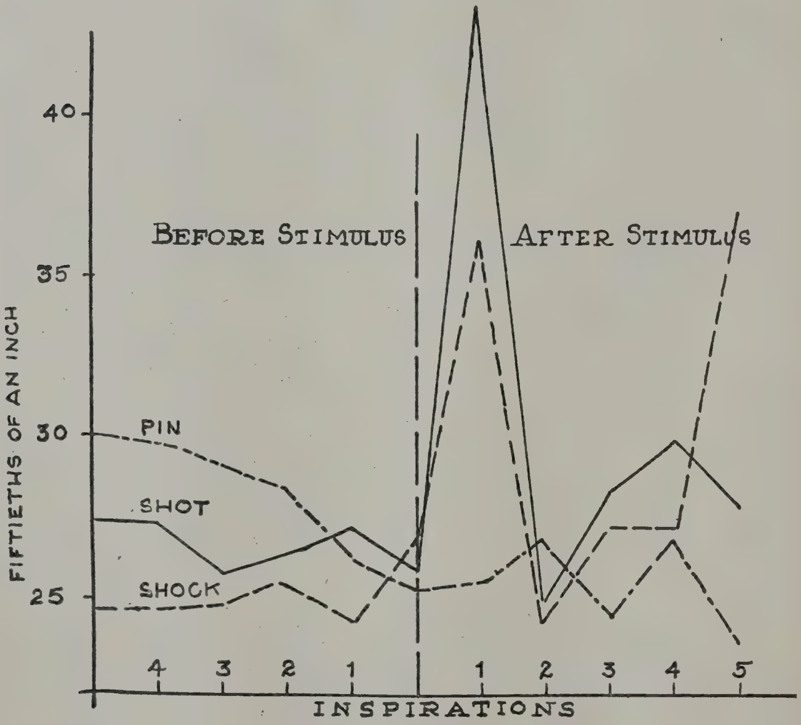


FIGURE 2

Composite curve showing the average amplitude of each inspiration derived from the means of 20 subjects for the five consecutive inspirations before and after each stimulation. Note the very deep inspiration following the stimulus, succeeded at once by a very shallow breath, and then the gradual increase in amplitude.

Blatz (2, p. 128) observed that if the fear stimulus occurred during expiration this phase was immediately transformed into inspiration. If, on the other hand, the stimulus occurred during inspiration, inspiration was continued. Our records confirm this observation.

Another variable of breathing was discovered in connection with the amplitude of inspiration. When amplitude of breathing increases, the aggregate oscillatory movements of the stylus recording respiration are of course increased. A comparison of these aggregate movements before and after the stimulus should reveal a considerable difference. We were fortunate in possessing an instrument with which the vertical distance traversed by the stylus in recording respiration might be measured. This apparatus may appropriately be called the Hull linear oscillometer.[7] It was designed to give a measure of the vertical distance traversed by a stylus tracing any irregularly oscillating line. It is especially adapted to trace the irregularities which occur in the respiratory tracings of our records.

For this index we determine the total vertical excursion of the respiratory tracing for a period of 40 seconds (in the sample tracing for a period of 25 seconds) before and after the instant of stimulation. The difference between the two, as expressed in units of the instrument, furnishes a convenient measure of the changes in breathing oscillations due to the stimulus. Since individual differences in respiration are marked, it is necessary to reduce the difference between the measurements to percentage of decrease or increase as a result of the stimulus. These values as measured for Subject 11 on the reproduced record, Plate 1, are as follows:

Before stimulus, 201° After the stimulus, 386°
Percentage increase of the vertical excursions = 92

It is apparent that such a marked difference in respiration as indicated above may be due to either one or two factors or to a combination of both. The subject may react to the stimulus with an increase in the rate of inspiration for the interval of time measured or with a slower rate and an increased depth of breathing. Either of these changes might result in an increase of the aggregate oscillatory movement of the stylus. That individuals differ greatly in their respiratory disturbances is revealed by an examination of the records.

---

[7] This instrument is described in detail in (3).

To reveal any quantitative differences which might characterize the three stimuli used, the percentages derived from measurements of the respiratory oscillations were averaged for the 20 subjects. The computations are shown in Table 4. It is to be observed that the characteristic changes for this index are in the form of an increase in oscillation as a result of a stimulus. There seems to be some indication of a quantitative difference between the several stimuli. The shock and the shot are clearly the most potent agents in producing changes of this index. The reactions resulting from the pin, as usual, are relatively slight. Indeed, a number of large negative values appear in this column, indicating in those cases an actual decrease in the up-and-down distance traversed by the stylus in those particular cases. The larger average computed for the shot stimulus is the result of the influence of the score of Subject 10, whose respiratory oscillations were excessive. Individual variations in the form of both greater and smaller oscillations are more pronounced as a result of the shock stimulus. Only one negative value appears in the

TABLE 4

Showing the differences between the length of the respiratory tracing before and after stimulations, expressed in percentage of the length before. Negative values denote a decrease in oscillation following the stimulus.

Subject	Stimuli		
	Shock	Pin	Shot
1	.60	.09	.38
2	—.32	—.10	.60
3	—.03	—.25	.06
4	.31	.03	.20
5	.08	.07	.14
6	.33	—.33	.23
7	.03	—.11	.29
8	.053	.13	.04
9	.13	.09	.24
10	.21	.29	1.05
11	.68	.32	.70
12	—.11	.01	.25
13	.60	.18	.42
14	.40	.14	.00
15	.43	.11	.23
16	.06	.09	.32
17	.28	.25	.12
18	.24	(missing)	—.09
19	.45	.07	.08
20	—.08	—.06	.51
Means	.247	.027	.288

shot column, that for Subject 18. Apparently a strong auditory stimulus of this type produces the greatest oscillation of the stylus. The writer is inclined to the view that the percentage increase in respiratory oscillation is one of the best indices of emotional reaction.

In addition to the indices in respiration which we have already considered, there still remains the duration of the total respiratory cycle. The lines ruled on the respiratory tracing of Plate 1 enabled us to measure and compare the duration of any period of inspiration or of expiration of the total cycle for any convenient period. The measurements of the duration of inspiration and expiration in the sample record are found in Columns 5 and 6, respectively, for the period previous to the stimulus, and in Columns 11 and 12, respectively, for the period following the stimulus, Table 3. The duration of the total cycle for the two periods is found in Columns 7 and 13, respectively.

That there is a distinct tendency for both the inspiratory and expiratory, and consequently the whole cycle, to be shortened by stimuli such as an electric shock and a shot is evident. In Table 5 are presented the percentage of the subjects who show a decrease in both inspiration and expiration and the total cycle. While the actual decrease as expressed in percentage of the pre-stimulus value is not large, a decrease exists for each of the three indices.

We might expect that variations in duration of respiration in the form of deviations from the mean duration would be pronounced. Mean deviations computed from the respiratory tracing of Subject 11, Plate 1, are found in Column 9 before stimuli and in Column 15 after stimuli, Table 3. In Table 6 we have tabulated the results of computations of this index which we have termed the *mean deviation of the breath cycle.*

It is apparent that the average tendency of the group used in this experiment was toward a greater variability of duration of respiration after the stimulations. The greatest individual variability appears with the shot, while the pin again shows itself to be a weak source of disturbance. The computation of the mean deviation appears here, as well as with the previous measurements, a valuable indicator of the emotional disturbance attendant on strong sensory stimuli.

Benussi (1) was among the first to point out the value of the inspiration-expiration ratio as an indication of emotional disturbance. This well-known value is obtained by dividing the time for inspiration (*I*) by the time for expiration plus the short pause which

TABLE 5

Showing the percentage of the 20 subjects who responded with a decrease in the duration of inspiration and expiration and of the total inspiration-expiraton cycle, together with the averages for the three indices expressed in percentage decrease from the pre-stimulus value for each of the three stimuli.

	Shock	Pin	Shot
Percentage of the subjects who show a decrease in the duration of inspiration	60	40	60
Percentage decrease from the pre-stimulus value	9	3.8	2.9
Percentage of the subjects who show decrease in duration of expiration	65	70	50
Percentage decrease from the pre-stimulus value	9.7	5.3	4.1
Percentage of the subjects who show decrease in duration of the total cycle	75	65	75
Percentage decrease from the pre-stimulus value	8	5	5.5

TABLE 6

Showing the percentage of the 20 subjects who responded with an increase in deviations from the mean duration of respiration, together with the percentage increase which the average index for the group as a whole shows for the period following the stimulus over the pre-stimulus period.

	Shock	Pin	Shot
Percentage of the subjects showing an increase in the *mean deviation of duration of respiration*	65	20	60
Average response to stimulus expressed in percentage gain over the pre-stimulus value	23	.9	73

usually occurs after expiration (E). With the aid of the lines drawn on Plate 1, we can compute the inspiration-expiration (I/E) ratio. Columns 8 and 14 of Table 3 show the I/E ratio as computed for Subject 11. The index here falls from unity as a result of the stimulus. That this subject deviates from the tendency of the group on this index is shown by the average value, which was 1.235, or greater than unity. This is an increase over the pre-stimulus value of 295%. This large gain was, however, due to the responses of only 55% of the subjects. An increase in the value of the I/E ratio cannot, therefore, be taken as typical of the whole group. A very slight average increase in the I/E ratio was found to result from the shot. In spite of the fact that 60% of the subjects showed slight increases in the ratio after the pin stimulus, the average value for that stimulus is very slightly decreased. Certainly the I/E ratio fails in the case of the pin and the shot as an indicator of emotional disturbance when such a disturbance exists as evidenced by a majority of the other indices.

That variations in the deviations from the mean I/E ratio exist was proved by employing an index which we have termed the *index of variability in the I/E ratio.* This index is shown in detail for Subject 11, in Columns 10 and 16 of Table 3. It is derived by first computing the extent to which the I/E ratio deviates from the average. The mean of these deviations yields our new index. We find that the shot is here the most disturbing stimulus, although, as we have seen, the shock produced a greater increase in the I/E ratio. Seventy per cent of the subjects show an increase in the variability of the I/E ratio. There was an average increase of 75% over the pre-stimulus value. Sixty-five per cent of the subjects show an increase as a result of the pin stimulus, but the increase is largely neutralized by the remainder of the group, there being an increase for this stimulus of only 13% of the pre-stimulus value. Only 55% of the group show an increase in variability as a result of the shock. The average increase was, however, 49% of the pre-stimulus value.

Although influenced by sampling errors, the *index of variability of the I/E ratio* is probably at least of equal value to the I/E ratio itself as a useful indicator of emotional disturbance.

In Table 7 we have summarized the most important disturbances in heart and respiratory activity as measured. The figures given include the actual difference in values of the various indices between the periods before and after the stimuli, the percentage gain or loss

TABLE 7

Index	Actual difference between periods before and after stimuli	Percentage gain (+) or loss (—) in value	Percentage of 20 subjects showing gain	Actual difference between periods before and after stimuli	Percentage gain (+) or loss (—) in value	Percentage of 20 subjects showing gain	Actual difference between periods before and after stimuli	Percentage gain (+) or loss (—) in value	Percentage of 20 subjects showing gain
Amplitude of pulse	.422	+41	75	.026	+21	60	.70	+100	90
Amplitude of breathing	3.155	+11.4	75	.455	+16	35	5.26	+ 20	80
Deviations in amplitude of breathing	2.528	+69	75	.018	+ .5	40	7.441	+225	90
Duration of inspiration	.302	+ 9	40	—.117	— 3.8	60	.628	— 2.9	40
Duration of expiration	2.145	+ 9.7	35	—.59	— 5.3	30	.395	— 4.1	50
Deviations in duration of total cycles	.3395	— 8	25	.1005	— 5	35	.654	+ 5.5	25
I/E ratio	.9225	+295	55	—.004	+12	60	.010	+32	50
Variations in the *I/E* ratio	.390	+49	55	.065	+13	65	.585	+73	70

over the pre-stimulus value, and the percentage of the 20 subjects who showed a gain as a result of the stimuli.

4) *Subjective Report.* It will be recalled that at the conclusion of the experiment each subject was requested to rate his subjective reactions relative to the intensity value of the stimuli. While the averaging of ratings can give only an approximation, it is interesting to observe the result of pooling the judgments of the 20 subjects for each of the three stimuli.

For the shock—2.5 i.e., moderately intense
For the pin—3.7, i.e., moderate to slight
For the shot—2.0, i.e., intense

We have observed throughout the objective treatment the greater stimulating value of the shot. Not only does it provoke an increase in rate and amplitude of the various indices considered, but the pin stimulus has throughout our results proved itself an impotent stimulus in producing marked disturbances. There appears then a certain correlation between the subjective estimation of the intensity value of the stimuli and the variability of the physiological rhythms we have been considering.

That previous experience tended to produce a form of conditioning to several of the stimuli is indicated by the reaction of several of the subjects. Three of the four who showed the greatest disturbance to the shot had been members of the track team and were extremely sensitive to the firing of the gun. Examination of their results reveals not only unusually large disturbances in the rate and amplitude of the pulse and of respiration but also marked fluctuations of the psyochogalvanscope. On the other hand, two of the subjects showed a remarkable lack of sensitivity when stimulated by the electric shock. It was later learned by the experimenter that these two men were students in electrical engineering and were rather tolerant of the electric shock from frequent stimulation of this kind in the electrical laboratory. Their characteristic remarks on later questioning in reference to the shock stimulus were similar to, "Oh, that wasn't so bad." "It takes a lot of juice to jar me," etc.

SUMMARY

1) Varying amounts of change in the skin resistance of the subjects, always in the form of a reduction, result from all of the stimuli used in this experiment.

2) Previous conditioning to various forms of the stimuli used appears in the form of a less variable skin resistance on excitation.

3) With the apparatus and the pressure adjustments here used the pulse wave immediately diminishes in amplitude and then increases above normal for several seconds after the shock and shot stimuli.

4) Pulse rate increases in most cases to the two stronger stimuli, the greatest increase being after the shot.

5) The duration of the total respiratory cycle reaches a maximum at the third cycle following the shock and shot stimuli. The increase in the amplitude of inspiration (depth of breathing) is in direct proportion to the order of intensity of the stimuli.

6) The inspiration-expiration ratio (I/E ratio) exceeds unity for the five respiratory cycles succeeding the shock (average for the 20 subjects).

7) The total up-and-down oscillations of the respiratory tracing are characteristically increased as the result of stronger sensory stimuli. Exceptions to this rule appear with weaker stimuli.

8) Deviations from the mean were in all cases greater after the stimulus wherever this index was computed.

a) Deviations of the I/E ratio were most extensive after the shot.

b) Deviations in the duration of the total breath cycles and in the depth of breathing were both greater after the shot.

c) Deviations from the mean amplitude of the pulse for a period of 25 seconds after the shock and the shot are nearly twice as great as during the same period before these stimuli.

9) The subjective grading of the intensity value of the three stimuli was in strict accord with the relative disturbance of the various physiological rhythms as objectively determined.

10) It is suggested that some of the indices of emotional reaction employed above, notably the *mean deviation of pulse amplitude* and the *mean deviation of depth of respiration* compare very favorably in value with those ordinarily used in investigating the emotions.

REFERENCES

1. BENUSSI, V. Die Atmungssymptome der Lüge. *Arch. f. d. ges. Psychol.*, 1914, **21**, 224-273.
2. BLATZ, W. E. Cardiac, respiratory, and electrical phenomena involved in the emotion of fear. *J. Exper. Psychol.*, 1925, **8**, 109-132.
3. HULL, C. L. An instrument for summating the oscillations of a lines. *J. Exper. Psychol.*, 1929, **12**, 359-361.

4. Smith, W. W. A note on the use of the psychogalvanic reflex. *Brit. J. Psychol.,* 1921, **12**, 282-288.

University of Wisconsin
Madison, Wisconsin

LES RÉACTIONS ÉMOTIONNELLES AUX STIMULI FORTS

(Résumé)

On a essayé dans cette expérience d'étudier très soigneusement de certaines réactions physiologiques à des stimuli forts dans le but de découvrir des changements significatifs.

Un sphygomomanomètre faisant des notations, un pneumographe, et le galvanomètre ont été employés pour faire des notations du pouls, de la respiration, et du réflexe de la peau sur le tambour enfumé. On a étudié les réactions de vingt sujets à un coup de révolver, à un choc électrique, et à un coup d'épingle comme notées par les instruments cités ci-dessus. Outre la computation de la vitesse et de l'amplitude du pouls et de la respiration on a trouvé utiles plusieurs nouveaux indices. On a fait une mesure soigneuse des distances horizontales entre les crêtes des ondes respiratoires. On a comparé une détermination de la distance moyenne et des écarts de la moyenne pour une période immédiatement précédant chaque stimulus à celle pour la période immédiatement après. On a traité d'une manière semblable l'amplitude du pouls et de l'inspiration. On a trouvé que les écarts de l'amplitude moyenne du pouls et de la respiration et de fréquence respiratoire varient indépendemment de la simple vitesse et de l'amplitude. Ces indices ont donné une haute corrélation avec l'étendue du réflexe de la peau et l'estimation subjective de l'intensité des stimuli. On les offre comme des indicateurs utiles du degré de perturbation émotionnelle.

J. E. Caster

GEFÜHLSREAKTIONEN AUF STARKE REIZE

(Referat)

Dieses Experiment war ein Versuch gewisse physiologische Reaktionen auf starke Reize genauer Prüfung zu unterwerfen, um bedeutende Veränderungen aufzudecken.

Zum Registrieren des Pulses, der Atmung und der Hautreflexe auf berusster Trommel benutzte man einen registrierenden Sphygmomesser, einen Pneumographen und das Galvanometer. Man untersuchte die Reaktionen von zwanzig Versuchspersonen auf einen Pistolenschuss, einen elektrischen Schlag und einen Nadelstich. Neben der Berechnung der Raten, der Puls- und Atmungsausschläge wurden auch einige neue Bestimmungsmerkmale als wertvoll befunden. Man mass die horizontalen Distanzen zwischen den Bergen der Atmungskurven. Die Bestimmungen der Mittel und Abweichungen vom Mittel unmittelbar vor dem Reiz wurden mit denen dem Reiz unmittelbar folgenden verglichen. Die Puls- und Atmungsausschläge wurden in gleicher Weise behandelt. Abweichungen vom mittleren Ausschlag des Pulses und der Atmung und der Atmungsfrequenz varierten unabhängig von den blossen Raten und Ausschlägen. Diese Bestimmungsmerkmale korrelierten in hohem Grade mit der Ausdehnung des Hautreflexes und der subjektiven Schätzung der Reizintensität. Man schlägt vor, sie als brauchbare Indikatoren fur den Grad der Gefühlsstörungen zu verwerten.

Caster

THE INFLUENCE OF LENGTH OF PROBLEM AND OF TRANSFER UPON RATIONAL LEARNING AND ITS RETENTION*[1]

From the University of Arkansas

JOHN A. MCGEOCH AND VERA J. OBERSCHELP

PROBLEM AND METHOD

It is a question whether, and to what extent, the solution of problems which permit and favor a rational attack show the same general characteristics as does the memorization of a fixed series such as a list of words or of syllables, in which no special form of rational organization is favored. The retention values of the solutions of rational problems are matters of equal importance. The Peterson Rational Learning Problem, first described in 1918 (3) and used by several investigators since that time (1), lends itself well to an attack upon certain phases of this inquiry. It has the advantage that it can be constructed in any desired length, that it is objectively scorable, that the scores can be treated in a manner strictly comparable to the handling of syllable and word lists, and that it presents, nevertheless, a situation which is not as fixed as the conventional serial learning, but one in which the limitation of the range of probability by each succeeding response presents to the subject a constantly changing situation. Within these limits fixation must occur, but it is a fixation within a framework of rational analysis, the extent and efficacy of which depend, of course, on the subject.

In this paper data are presented upon the influence of two factors upon rational solutions and their retention: (*a*) length of problem and (*b*) transfer. Twenty-four college girls[2] learned 6-, 12-, and 18-letter rational learning problems to a criterion of two successive perfect trials at a single sitting and relearned them to the same criterion after an interval of seven days. The problems were learned

---

*Received for publication by Carl Murchison of the Editorial Board, February 3, 1930.

[1] Research paper written at University of Arkansas.

[2] The subjects for this experiment were furnished to Miss Oberschelp by Professor Louise Holcomb of Wilson College, to whom we wish to express our thanks.

and relearned one at a time. That is, the first problem was learned and, after a week, was relearned; then an interval of one or two days elapsed between this relearning and the learning of the second problem, and again between the relearning of this problem and the learning of the third. Any significant interference between problems was thus avoided.

Since 36 letters are required for the three problems, both the English and the Greek alphabets were employed. One problem of each of the three lengths was constructed from each alphabet, with the two sets of letters at each length from corresponding positions in the two alphabets, and with identical numbers assigned to each set. This leaves the character of the letters as the only variable. The order in which the six conditions (two problems at each of three lengths) were gone through was counterbalanced to give the conditions equal average position.

At the beginning of each sitting the subject was given a typewritten sheet of instructions and asked to read them through twice and to ask any questions necessary to clarify their meaning. The instructions and the subsequent procedure were copied from Peterson's paper, with the addition of the procedure in relearning. At the close of the learning period each subject was asked to return after a week for the relearning and was cautioned with great care against rehearsing the problem in any way and against talking about it with anyone. When each subject returned for relearning, she was asked, in detailed questions, whether she had followed these instructions, and only those who had followed them were included in the study. The subjects were well known to the experimenter and there is good reason to believe that the influence of intervening rehearsal and discussion was at a minimum. The second addition to Peterson's technique was the requirement of a supplementary relearning of the problem with the letter-number pairs presented in a changed order. In another experiment one of the writers observed that the subjects sometimes learned the numbers as a series, fitting them together without reference to their paired letters. Rational analysis may still play a part when this method is used, but since the memorization aspect of the problem is designed to be of the paired-associates variety, serial memorization is undesirable. Consequently, in this experiment, the subjects were instructed to associate the letter with its number, and the supplementary relearning with the stimulus letters in a changed order, following immediately after relearning

TABLE 1*

MEAN SCORES AT EACH LENGTH

Problem	Pairs presented	Time (min.)	Logical errors	Perseverative errors	Unclassified errors
6-letter	14.75	2.85	4.58	.166	10.79
12-letter	82.00	20.44	57.33	13.66	86.45
18-letter	111.75	30.13	92.62	22.75	161.87

*The general consistency of the results implies the validity of their general tendencies. It is hardly necessary, therefore, to cumber the tables with the sigmas of the means and differences.

with the same order as in learning, was used to check the extent to which serial memorization had been employed. Presumably any pronounced inability to adjust to the changed order should indicate that the fixation of the solution in paired-associates form had not been followed, while an easy adjustment to the new order should indicate that the instructions had been followed. A slight difficulty of adjustment might be expected, but not one of any great extent.

The scoring was in terms of (*a*) number of trials; (*b*) time per trial and total time; (*c*) logical errors—guessing a number already used for an earlier letter of the series; (*d*) perseverative errors—repeating a wrong guess while reacting to a single letter; and (*e*) unclassified errors—all other wrong guesses. Each of these scores has been obtained for learning, relearning with the same order, and relearning with the changed order. The data for trials cannot, of course, be treated directly in terms of each problem as a unit, since the problems are not equal. Mean pairs presented, mean time, and mean errors per problem are, however, comparable scores. The time, trial, and error scores have also been converted into amounts per pair (letter and number), since pairs may be treated as equal units independently of the length of the problem in which they appear. The details of the manner in which these scores have been used can be described more clearly in direct connection with the separate phases of the problem.[3]

---

[3] The results from the English and Greek alphabets at each length have been combined in all of the data reported. Each score has been computed separately for the two alphabets, however, to discover if combination is justifiable. While differences appear, the general tendencies are the same, and the differences are neither large enough nor consistent enough to warrant separate treatment of the problems from the two alphabets.

THE INFLUENCE OF LENGTH OF PROBLEM UPON DIFFICULTY OF LEARNING

The mean scores up to, but not including, the criterial trials are given in Table 1.

If each length of problem were strictly proportional to each other length in difficulty, we should expect mean pairs presented, mean time per problem, and mean errors per problem to be in the ratio 1:2:3. Actually, however, the 12- and 18- letter problems are more difficult than the 6-letter problem to a degree which is markedly disproportionate to their respective lengths. (See Table 2). This is true for all scores, though in varying degree. The disproportion between the 6-letter and the 12- and 18-letter problems decreases in the order: perseverative errors, logical errors, unclassified errors, mean time, and mean pairs presented. The ratios between the 12- and 18-letter problems (the figures in parentheses after the 18-letter problem in Table 2) are much more nearly in proportion to the lengths of the problems. If difficulty were in proportion to length, we should expect the 18-letter problem to have scores 1.5 larger than those for the 12-letter problem. The experimentally obtained ratios between the two problems are slightly less than this for mean pairs presented and for mean time and somewhat greater than this for all error scores.

TABLE 2

RATIOS OF DIFFICULTY

Problem	Pairs presented	Time (min.)	Logical errors	Perseverative errors	Unclassified errors
6-letter	1	1	1	1	1
12-letter	5.55	7.17	12.51	82.28	8.01
18-letter	7.57 (1.36)	10.57 (1.47)	20.22 (1.61)	137.04 (1.66)	15.00 (1.87)

It must be concluded, therefore, from these data that 12- and 18-letter rational learning problems are more difficult to learn than a 6-letter problem to a degree which is markedly disproportionate to the lengths of the respective problems. The increase in length from 12 to 18 letters is accompanied by an increase in difficulty which is more nearly in proportion to the increase in length.

Further evidence for the same conclusion is found when mean seconds per pair and mean errors per pair are examined. If the three problems are difficult in proportion to their lengths, we should expect mean time and mean errors per pair to be identical for each problem. They could also be identical if the disproportion in difficulty were to affect only the mean number of pairs presented. We already know that the problems increase in difficulty faster than they increase in length, and we shall not expect mean time and errors per pair to be identical. Further, the greater disparity in the ratios for time and errors than in that for pairs presented would also prevent us from assuming that the disproportionate increase in difficulty holds only for pairs presented and for time and errors only through the former score. A more exact picture of the situation can be obtained, however, from Table 3, which presents the mean time and errors per pair and the ratios of the scores at the different lengths.

It is clear that increasing length of problem produces, as well as a disproportionate increase in the number of pairs which must be presented before learning occurs, a disproportionate increase in time per pair and in errors per pair.

TABLE 3

MEAN TIME AND ERRORS PER PAIR: AND THEIR RATIOS OF DIFFICULTY

	Time		Errors					
Problem	(sec.)	Ratio	Log.*	Ratio	Per.†	Ratio	Unc.‡	Ratio
6-letter	11.63	1	.310	1	.011	1	.731	1
12-letter	14.95	1.28	.699	2.25	.166	15.09	1.054	1.44
18-letter	16.18	1.39	.828	2.67	.203	18.45	1.450	1.98

**Log.* = logical errors.
†*Per.* = perseverative errors.
‡*Unc.* = unclassified errors.

That is, on the average, each unit of a single pair must not only be presented a disproportionately greater number of times as the length of the problem increases, but each pair requires an increasing time for mastery and an increasingly greater number of errors. The disproportion is greater between the 6- and 12-letter problems than between the 12- and 18-letter problems. The ratios of the two latter are not shown in the table, but can be computed readily. The disproportion runs in all cases in decreasing amount thus: perseverative errors, logical errors, unclassified errors, and time; with the single exception that between the 12- and 18-letter problems the disproportion is greater for unclassified than for perseverative errors.

The Influence of Length of Problem upon Retention

We may measure the retention values of the three problems in four different ways which will be described at this point. (*a*) The mean time taken and the mean errors made on the first relearning trial, and the mean time and errors per pair on this trial, will be designated *recall scores.* (*b*) These recall scores have been divided by the corresponding scores on the initial learning trial, and subtracted from 100, to give *saving-score*–1.[4] The records of the first learning trial have been used for this purpose since they represent the nearest approach to complete lack of knowledge of the problem, such as would be represented, for example, by the total number of syllables in a list none of which is known prior to the learning period. The saving score computed from the data of the initial learning and relearning trials states the extent to which the subject is able to handle the problem better after a week of no practice than at the first attempt prior to practice. (*c*) The mean number of pairs necessary to present for relearning, and the mean time and errors, together with the mean time and errors per pair, constitute *relearning scores.* (*d*) These scores divided by the corresponding records for learning, and subtracted from 100, provide *saving-score–2.* In computing this and the relearning score the data for the two criterial relearning trials have been omitted, but in computing saving-score–2 the records for the criterial learning trials have been added to the number of pairs presented and the time, as given in Tables 1 and 3.

In order to avoid cumbersomeness in the tables which follow, the ratios of the scores at the three different lengths of problem have been omitted. They can be determined roughly by inspection, or exactly by a few minutes' computation. The recall scores and the saving-scores–1 will be found in Table 4.

The mean recall scores per problem increase out of proportion to the increase in length, but the disproportion is less marked than in learning. The saving scores computed from mean scores per problem mirror this tendency to a levelling in retention of differences found in learning, although there is no exception to the generalization that the problems harder to learn are the better retained. In both recall

---

[4] The formula used is: Saving $= 100 - \frac{\text{First relearning trial}}{\text{First learning trial}} \times 100$. In finding *saving-score–2* the total learning records and the relearning records up to the criterial trials have been used instead of initial trials.

TABLE 4

Recall Scores and Saving-Scores–1

	Mean scores per problem				Mean scores per pair			
	Time	Errors			Time	Errors		
Problem	(min.)	Log.	Per.	Unc.	(sec.)	Log.	Per.	Unc.
				Recall Scores				
6-letter	.94	1.25	.166	2.37	9.41	.20	.02	.39
12-letter	2.60	3.95	1.00	10.45	13.03	.33	.05	.87
18-letter	3.62	8.70	1.04	13.45	12.09	.48	.05	.75
				Saving-Scores–1				
6-letter	12.96	30.17	–315.00	70.66	15.31	35.48	–233.33	61.76
12-letter	47.87	76.17	77.97	80.41	47.85	76.08	78.37	67.77
18-letter	57.75	78.53	87.76	88.36	57.68	78.66	89.36	79.67

scores and saving-scores–1 the disproportion is greatest when the 6-letter problem is used as a base. In general the 12- and 18-letter problems tend to approach each other much more nearly in retention than in learning. In every case save mean logical errors per problem the 18-letter problem is less than 1.5 times better retained than the 12-letter problem.

Mean time and errors per pair are likewise greater for the 12- than for the 6-letter problems in recall, while the 12- and 18-letter problems tend to be close together. Saving scores for mean time and errors per pair evidence the same tendency as do those for mean time and errors per problem. In both cases the high percentages of loss in perseverative errors is of little significance, since, presumably by chance, a very small number of these errors was made on the initial learning trial.

TABLE 5

Relearning Scores and Saving-Scores–2

		Mean scores per problem				Mean scores per pair			
	Mean pairs	Time	Errors			Time	Errors		
Problem	presented	(min.)	Log.	Per.	Unc.	(sec.)	Log.	Per.	Unc.
				Relearning scores					
6-letter	7.00	1.16	1.62	.25	3.08	9.96	.232	.035	.440
12-letter	33.00	4.89	6.25	1.54	15.54	8.93	.189	.046	.471
18-letter	38.25	6.28	13.70	1.79	19.45	9.85	.358	.046	.562
				Saving-Scores–2					
6-letter	73.83	67.50	64.62	–50.60	71.45	–24.34	–35.66	43.54	–09.18
12-letter	68.86	77.61	89.09	88.72	82.02	27.80	65.06	64.06	42.33
18-letter	74.11	80.40	85.20	92.13	87.98	24.28	42.90	70.12	48.67

Relearning and saving-scores–2 are given in Table 5. The mean number of pairs required to be presented for relearning the three problems stand to each other in only a slightly less disproportionate relation than in learning. The ratios between the mean time and error scores per problem are, likewise, far from the ratio 1:2:3, but the disproportion is much less than in learning. This tendency toward a levelling of the differences between the three problems becomes more pronounced in the saving scores. The problems differ relatively little in saving in pairs presented, and the saving in time and errors is in a ratio less than that of the lengths of the problems. Only in pairs presented and in logical errors, however, do the saving scores fail to increase uniformly with increasing length of problem.

Mean time per pair and mean logical errors per pair in relearning vary, but fail to show any uniform relation to length, while mean perseverative and unclassified errors per pair tend to increase slightly with length. The saving in terms of mean scores per pair is distinctly less than in terms of mean scores per problem. With the 6-letter problem there is a considerable loss for all scores except perseverative errors. Mean time and mean logical errors per pair show a saving which increases from the 6- to the 12-letter problem and decreases to the 18-letter problem, and saving in perseverative and unclassified errors increases with increasing length of problem. The same general levelling of the differences between the problems is evident here.

It will have been observed that the retention values of the rational learning problems, particularly of the two longer, are relatively high. If we omit the 6-letter problem, which is retained much less well throughout than the others, the range of saving-scores–1 is from 47.85 to 89.36, with 11 out of 16 scores coming above 70. The range of saving-scores–2 is from 24.28 to 92.13, with 10 out of 18 scores above 70. If we consider only saving in mean scores per problem, all but one score is above 70. The tables show clearly that time and pairs presented yield less saving than do errors, and that mean scores per pair, especially with saving-scores–2, give less saving than do mean scores per problem. Saving scores as high as these after an interval of a week are seldom found with the usual type of learning problem.

We may compare these values with those obtained from three mazes of differing degrees of difficulty and from lists of 8, 12, and 16 nonsense syllables. A paper reporting the retention values of the mazes and syllables has recently been published (2). The conditions of the two experiments are closely similar, and retention has been measured after an interval of a week in both cases. In the rational learning experiment all the subjects were girls, while in the maze-nonsense-syllable experiment only half were girls; but it is not known that, and it is doubtful if, the retention of rational learning problems yields any striking sex differences. The chief difference between the two experiments lies in the fact that the learning and relearning criteria were for rational learning two perfect trials in succession, and for the maze–nonsense-syllable experiment one and three, respectively. Thus the balance is in favor of the rational problems in learning and against them in relearning. The precise influence of this difference is unknown, but with it in mind the scores may be compared. The table of saving scores for the mazes and syllables is reproduced here as Table 6.

TABLE 6

Maze and Nonsense Syllable Saving Scores

	Trials		Time		Errors	
Maze and list	Maze	N.S.	Maze	N.S.	Maze	N.S.
Easy (8-syllable)	–18.30	43.19	31.00	43.19	60.46	58.64
Medium (12-syllable)	01.05	34.91	48.62	34.91	76.51	66.72
Hard (16-syllable)	–11.38	49.35	55.56	49.35	77.00	74.39

Saving-scores–2 are the ones which are comparable to the data of this table. In terms of mean pairs presented, which are the direct analogue of trials, the rational learning problems of all three lengths are retained far better than are the mazes. The difference between rational learning and nonsense syllables is less, but still uniform and considerable. The difference between the mean saving scores is 29.78. In saving in time, the difference between the means for rational learning and mazes is 30.11, and between those for rational learning and the syllables 32.69. The mean saving in errors is for the mazes 71.32 and for the syllables 66.58, while for rational learning it is as follows: logical errors, 79.63; perseverative errors, including the unreliable loss with the 6-letter problem, which we need not consider seriously, 43.42; and unclassified errors, 80.48.

The conclusion is well supported that by all comparable measures the rational problems are better retained than are either mazes or nonsense syllables. The better retention of rational learning than of either mazes of nonsense syllables is probably a function of the greater unity and meaningfulness which, as a result of the favored rational and analytic attack, the subject is able to create in the problem.

It will be remembered that, when each problem had been relearned to a criterion of two perfect trials, its items were at once presented to the subject in a changed order until she had reached two successive perfect trials with this order. One could scarcely expect the subjects to continue their perfect adaptation to the problem under the new conditions, but if they had followed the instructions to learn the problem as a series of paired associates, the period of adaptation to the new order of letters should be short.

Some subjects were able to "shift" readily and to adapt immediately to the new order. The numbers of such subjects were 13, 9, and 7 for the three problems, as against 6, 1, and 1 who gave errorless trials at the first relearning attempt with the same order. Others required from one to six trials before the criterion was attained. In terms of saving in pairs presented, which is the basic score, using learning plus relearning with the same order as the denominator, the savings are 85.24, 87.11, and 86.48. The savings in time and errors per problem are of the same order of magnitude. Saving in time and errors per pair is much less and in a few cases there is even a loss. These latter facts indicate a short and difficult initial period of adjustment, during which errors were frequent and when a comparatively long time was taken for each choice.

On the whole, it seems that we may fairly assume that the instructions were carried out sufficiently to avoid any vitiation of the conclusion that we have been dealing with the fixation of rationally selected paired associates and not with the acquisition of a serial order, although it is certain that some effects of serial order were present. It would be difficult and, for our purposes, unnecessary to say to what exact degree this is the case. The fact that a large percentage of the subjects required some little time for the acquisition of the new order indicates the presence of some serial connections and suggests the need for care in controlling the method by which subjects attack the fixation of rational learning problems.

Transfer (Practice Effects)

The influence exercised upon speed of learning by practice on other problems of different lengths may be shown by grouping together the scores of all the subjects coming at each of the three possible positions in the practice order. In the learning of each problem there are eight subjects who attack that problem first, with no prior experience, eight who attack it after having learned one other problem, and eight who attack it after experience with two others. It would be possible to divide the data in order to discover the effect of particular antecedent problems, also, but the very small number of subjects in each order would render such a division worthless. We shall confine ourselves, therefore, to inquiring the effect of no preceding practice compared with that of practice upon one and upon two preceding problems. Since there are but eight subjects in each sub-group, even here, we must depend upon the consistency of the tendencies to establish their validity.

In Table 7 the numbers I, II, and III refer to the position in the practice order. The scores are for learning up to the criterial trials. It is clear that, with the 6- and 12-letter problems, ease of learning is greatly increased by practice at other problems and that practice on two problems yields a distinctly greater increase than does practice on one. This increase from the first to the third positions in the practice order is uniform in tendency for all scores. It differs somewhat in amount, being least of all for unclassified errors and greatest for perseverative errors. There is, of course, small opportunity for practice to affect unclassified errors, while a complete eradication of the other types of error would be possible. There is no doubt that preceding practice at learning 12- and 18-letter problems transfers markedly to the learning of the 6-letter problem, nor that practice at 6- and 18-letter problems transfers considerably to the 12-letter problem.

TABLE 7

Scores for Each Problem at Each Position in the Practice Order

Position	Pairs presented	Time (min.)	Errors Log.	Per.	Unc.
			6-letter problem		
I	21.00	4.85	7.62	.50	13.62
II	13.50	2.15	3.83	.00	9.87
III	9.75	1.46	2.12	.00	8.87
			12-letter problem		
I	112.50	32.32	104.62	24.25	121.75
II	75.00	18.33	44.62	14.75	74.25
III	58.50	10.66	22.75	2.00	63.37
			18-letter problem		
I	121.50	29.37	117.12	21.62	170.50
II	108.00	31.62	86.00	24.62	172.12
III	105.75	29.38	74.75	22.00	143.00

Practice at 6- and 12-letter problems helps the learning of the 18-letter problem relatively little, however. Pairs presented and logical errors show a uniformly positive practice effect, while time and perseverative errors show no uniform tendency. There are at least two possible explanations for the absence of a practice effect upon the 18-letter problem. It may be that this problem, on account of its greater length and difficulty, requires a mode of attack sufficiently different from that required by the shorter problems to render practice on them of little assistance. It seems a more likely hypothesis, however, that transfer is a function of the amount of preceding practice, measured, not in terms of problems, but of pairs presented, time, and errors. There has been relatively less practice prior to the 18-letter problem at its two transfer positions than prior to the corresponding positions with the two other problems, where the 18-letter problem, which requires a number of pairs presented, and time and error scores, which are more than the sum of the corresponding scores required by the two other problems must have been one of the two preceding problems. Hence we should expect less transfer to it from the shorter problems than from it to them. This hypothesis is borne out suggestively by an examination of the amount of transfer from each of the shorter problems to each other. To the conclusion that practice at rational learning problems markedly increases the speed of their acquisition, we may add the tentative statement that the amount of such increase is a function of the amount of the preceding practice measured in actual learning scores which are, in turn, presumably indicators of the learning effort expended. The same general results appear when initial trials, instead of all trials, are examined.

When we turn from learning to retention, the beneficial effects of prior practice disappear. Whether we consider recall or relearning scores, the second and third positions in the practice order not only fail to maintain the advantage which accrued to them in learning, but they are, if anything, worse than the initial position in the order. The general tendency of the results upon the retention values of the different practice positions can be shown sufficiently by the saving-scores–2, omitting the lengthy tables of recall and saving-scores–1, all of which present the same picture (Table 8).

The scores for the 6- and 12-letter problems show a clear tendency for the second and third positions to be at a clear disadvantage in comparison with the first position. This tendency appears less dis-

TABLE 8

Saving-Scores–2 at the Different Practice Positions

Position	Pairs presented	Time (min.)	Errors		
			Log.	Per.	Unc.
			6-letter problem		
I	77.27	83.30	90.15	50.00	80.76
II	70.58	54.72	34.72	Inf.	65.85
III	72.41	59.55	23.58	Inf.	63.35
			12-letter problem		
I	75.82	84.69	94.86	94.35	90.04
II	62.42	73.69	88.42	87.32	77.27
III	65.24	63.90	70.32	31.50	72.19
			18-letter problem		
I	72.85	83.29	81.08	93.06	84.60
II	78.12	82.03	88.37	91.87	91.12
III	71.42	75.80	88.13	91.50	88.28

tinctly and less consistently with the 18-letter problem. This result that the positions most favored in learning are least favored in retention is a special case of what seems to be a more general principle of experimentation on learning, viz., that there often exists at least an apparent inverse relation between the efficacy of learning methods or learning conditions and the percentage of retention which they yield. It is experimental fact that the second and third positions were optimal in learning and are at a disadvantage in retention, and it is a plausible hypothesis that here, as in other parallel cases, the smaller amount of effort required for learning to the criterion led to a lessened retentiveness compared with that yielded by the greater (or at least longer) expenditure of learning effort demanded by the early positions. This assumes a positive relation between effort expended in learning and amount retained, a relation which it is beyond the scope of this paper to do more than suggest, but it seems a highly plausible assumption. The facts regarding the retention values of lists and problems of different lengths, e.g., bear it out. We have here a case in which the factors producing a transfer from the learning of one problem to the learning of another do not produce a like effect on retention, even when the retention is tested before the next problem has been learned. It seems as if compensating or antagonistic factors are at work, and it is suggested that learning effort is one of the most important. Thus, transfer factors render a problem more easily learned, but this saving in effort has a negative influence upon retention. The seeming failure of the results from the 18-letter problem to up-hold this generalization as unequivocally as do those from the 6- and 12-letter problems actually but adds support to it, inasmuch as the learning values at the three positions were, for the 18-letter problem, more nearly equal.[5]

SUMMARY

Three Peterson Rational Learning Problems, 6, 12, and 18 letters in length have been learned by 24 college girls and relearned after an interval of a week. The results have been examined in order to determine the influence of length of problem and of practice upon rational learning and its retention. The general conclusions follow:

1) Rational learning problems of 12 and 18 letters are more difficult to learn than are 6-letter problems to a degree which is markedly disproportionate to their respective lengths. An increase in length from 12 to 18 letters is accompanied by an increase in difficulty which is more nearly in proportion to the increase in length. These statements are supported by both mean scores per problem and mean scores per pair.

2) In terms of recall scores and saving-scores–1, there is no exception to the finding that the problems harder to learn are the better retained. The disproportion between the 6-letter problem and the two others is great. The 12- and 18-letter problems approach each other more nearly in retention than in learning. These conclusions hold for mean scores per problem and for mean scores per pair.

3) When relearning scores per problem are considered, the mean number of pairs required to be presented for relearning each problem stand to each other in only slightly less disproportionate relation than in learning. The ratios between mean time and error scores per problem are, likewise, not in proportion to length, but the disproportion is much less than in learning. Mean scores per pair fluctuate and only with unclassified errors show a consistent tendency to increase uniformly with length. When saving-scores–2 are

---

[5] Garrison (1) has reported no transfer effects from one 8-letter problem to another, when transfer is sought in terms of mean centile scores. Certain practice effects do appear when the learning curves are plotted. It should be noted, however, that the results of his experiment and of the present one are hardly comparable, with respect to transfer, on account of the divergence in procedure. Garrison studied transfer from one problem to another of equal length, and, also, when the subjects had had a large amount of individual practice aside from the taking of the rational learning test studied.

computed, there is still a definite tendency, to which mean pairs presented, logical errors per problem, and mean time and logical errors per pair are slight exceptions, to an increase in saving with increasing length of problem.

4) All measures of retention thus support two conclusions: (*a*) an increasing length of rational learning problems is accompanied by an increase in retention which is not, however, in proportion to the lengths of the problems; and (*b*) there occurs in retention a clear tendency toward a levelling of the differences which, in learning, appear between the problems.

5) Rational learning solutions are retained considerably better than are either mazes or nonsense syllables, when examples of the latter two materials are chosen from a wide range of difficulty.

6) The ease of learning 6- and 12-letter problems is greatly increased by practice at other problems; and practice at two preceding problems yields a distinctly greater increase in ease than does practice at one. Transfer to the 18-letter problem occurs, but not consistently nor to any large degree. This is possibly explicable by the asumption that transfer is a function of the amount of preceding practice measured in terms of pairs presented, time and errors. We should, then, expect less transfer from the shorter to the longer problems than vice versa.

7) Rational learning problems present certain of the same characteristics as does the memorization of a fixed series: (*a*) the difficulty of solution increases disproportionately to increases in length; (*b*) the longer problems (those harder to master) are retained better than the shorter; and (*c*) marked transfer effects appear from one problem to another.

References

1. Garrison, K. C. An analytic study of rational learning. *Peabody Coll. for Teach. Contrib. Educ.*, 1928, No. 44, Pp. 52. (Contains a summary of the work on rational learning.)
2. McGeoch, J. A., & Melton, A. W. The comparative retention values of maze habits and of nonsense syllables. *J. Exper. Psychol.*, 1929, **12**, 392-414.
3. Peterson, J. Experiments in rational learning. *Psychol. Rev.*, 1918, **25**, 443-467.

University of Missouri
Columbia, Missouri

New York State Reformatory for Women
Bedford Hills, New York

L'INFLUENCE DE LA LONGUEUR DU PROBLÈME ET DU TRANSFERT SUR L'APPRENTISSAGE RATIONNEL ET SA RÉTENTION

(Résumé)

Vingt-quatre étudiantes universitaires ont appris trois Problèmes Peterson de l'Apprentissage Rationnel, d'une longueur de 6, de 12 et de 18 lettres, et elles les ont appris de nouveau huit jours après. On a examiné les résultats pour déterminer l'influence de la longueur du problème et de la pratique sur l'apprentissage rationnel et sa rétention. Les problèmes de l'apprentissage rationnel de 12 et de 18 lettres sont plus difficiles à apprendre que les problèmes de 6 lettres à un degré fort disproportionné à leur longueur. L'accroissement de la longueur de 12 à 18 lettres est accompagné d'un accroissement de la difficulté lequel est plus proportionné à l'accroissement de la longueur. Toutes les mesures de la rétention soutiennent les conclusions qu'un accroissement de la longueur est accompagné d'un accroissement de la rétention lequel n'est pas cependant proportionné aux longueurs des problèmes comparés, et que dans la rétention il se trouve une tendance marquée vers un nivellement des différences qui se trouvent entre les problèmes dans l'apprentissage. Les solutions de l'apprentissage rationnel sont beaucoup mieux retenues que les habitudes du labyrinthe ou les syllabes non-sens. La pratique d'autres problèmes rend beaucoup plus facile l'apprentissage des problèmes de 6 et de 12 lettres; et la pratique de deux problèmes précédants donne un plus grand accroissement que la pratique d'un seul problème. Il y a un transfert au problème de 18 lettres, mais ce n'est ni constant ni très marqué. On pourrati expliquer ces résultats par la supposition que le transfert est une fonction de la quantité de pratique précédante mesurée en termes des résultats eux-mêmes. Les problèmes de l'apprentissage rationnel présentent beaucoup des mêmes caractéristiques présentées par la mémorisation des séries fixes de choses.

McGEOCH ET OBERSCHELP

DER EINFLUSS DER AUFGABENLÄNGE UND DER ÜBERTRAGUNG AUFS RATIONALE LERNEN UND DESSEN BEHALTEN IM GEDÄCHTNIS

(Referat)

Vierundzwanzig Collegemädchen lernten drei rationale Lernaufgaben von Petersen, bestehend aus 6, 12 und 18 Buchstaben, und wiederholten das Erlernen nach einer Woche. Man prüfte die Ergebnisse, um den Einfluss der Aufgabenlänge und der Übung auf das rationale Lernen und dessen Behalten im Gedächtnis zu bestimmen. Das rationale Erlernen von 12 bis 18 Buchstaben langen Aufgaben war schwerer als das von 6 Buchstabenaufgaben, und zwar in einem Masse, das die Proportionen der Länge deutlich übertrifft. Die Zunahme der Länge von 12 auf 18 Buchstaben ist von einer Zunahme der Schwierigkeit begleitet, die der Zunahme der Länge nahezu proportioan ist. Alle Masse des Behaltens bestätigen die Schlussfolgerung, dass die zunehmende Länge das Behalten unterstützt, aber nicht in proportionalem Verhältnis; dass das Behalten ein Ausgleichen der Unterschiede hervorruft, die zwischen den Aufgaben bestehen. Rationale Lernaufgaben werden besser im Gedächtnis behalten als Labyrintgewöhnungen und sinnlose Silben. Das Erlernen von 6 und 12 Buchstaben langen

Aufgaben wird durch Übung an andern Aufgaben sehr erleichtert; und zwar erleichtert die Übung mit zwei Problemen das Erlernen mehr als die Übung mit einer Aufgabe. Es kanneine Übertragung der Fähigkeit auf 18 Buchstaben lange Aufgaben festgestellt werden, sie ist aber weder konstant noch bedeutend. Diese Ergebnisse können möglicherweise durch die Annahme erklärt werden, dass Übertragung eine Funktion der verhergehenden Übung ist, und zwar in wirklichen Raten gemessen. Das Erlernen rationaler Aufgaben zeigt in mancher Hinsicht Übereinstimmung mit dem Memorieren von bestimmten Sachgruppen.

McGeoch und Oberschelp

THE INFLUENCE OF SOME AFFECTIVE FACTORS UPON RECALL*

From the Educational Psychology Laboratory of the University of Texas

Helen Lois Koch

The Problem and Its Historical Setting

It is the purpose of the study to investigate by way of a somewhat new approach one aspect of the problem of the relationship between the recall of experiences and their affective colorings. The colorings whose influences are to be probed are the pleasant and the unpleasant.

The general question is one with which psychologists have long concerned themselves. Within its purview many have placed the problem of selection in learning. The concept of repression also represents an effort to describe one phase of the relationship. Although attacks on the general subject have not been in the main experimental, even experimental approaches abound. The following enumeration of some of the criteria or methods used in attempting to determine the relative mnemonic power of pleasant and unpleasant experiences will attest our claim of popularity for the problem: the relative ease or frequency of the recall of pleasant and unpleasant memories as revealed by the questionnaire returns from lay groups (4); the relative numbers of the pleasantly- and unpleasantly-toned experiences among 100 miscellaneous ones spontaneously recalled or among the remembered events of a given day (11, 15); the serial position relative to each other of those class members later reported as most unpleasant and most pleasant in the lists of objects or events of certain kinds made by adults (19); the relative time involved in the verbal free-association type of response to words symbolic of things disagreeable, agreeable, or neutral (28); the relative time required for reading, writing, or learning lists of words of allegedly different affective significance, as well as the degree of retention of such lists after 24 hours (1, 3, 10, 20); the periods elapsing between the presentation of a so-called neutral stim-

---

*Recommended by Walter S. Hunter and received for publication by Carl Murchison of the Editorial Board, May 29, 1929.

ulus word and the arousal of pleasant or of unpleasant associations (26); the relative numbers of pleasant and unpleasant words written in equivalent units of time (10); the retention of items memorized when the learning period is followed immediately by a disagreeable experience such as the start stimulated by the unexpected discharge of a gun (26); the retroactive inhibition effects measured in the traditional way of lists of agreeable and disagreeable words (27); the relative retention values of colors and designs of known preference values (7, 26); the extent of recall of the names of affectively-toned odors or of arbitrary symbols which had previously been simultaneously presented to subjects with actual odors of varying degrees of agreeableness (8, 1); the incidence of unpleasant and pleasant memories among the very earliest ones which can be reinstated by adults (9); the memory values as determined by frequent tests of recall over a period of about a month of words associated with psychogalvanic responses of various degrees and with various combinations of so-called complex indicators.

Most of the investigations have revealed a readier arousal of pleasant memories than unpleasant, whereas Gordon's (7, 8, 9), Chaney and Lauer's (3), and Wohlgemuth's (31) studies, two of which at least are probably among the best controlled in the field, detected no difference in the extent of recall of satisfying and dissatisfying experiences.

The limitations of many of the techniques in the foregoing enumeration are obvious. Particularly conspicuous is the failure of some to disentangle the influence of the relative frequency of pertinent pleasant and unpleasant experiences from that of affective quality itself. One may remember, for instance, more pleasant occurrences than unpleasant, or write in a given time more words possessed of the former affective tone than the latter, or even give a verbal response more readily to agreeable than to disagreeable stimuli merely because he has had more pleasant than unpleasant experiences from which to draw his memories. This criticism becomes trenchant in the light of Flugel's (5) study which revealed that a larger proportion of our time each day is occupied in a fashion pleasing to us than otherwise. His subjects, recording at frequent intervals throughout their waking hours for 30 days the percentages of the time tinged with agreeable and disagreeable feelings, found that the former predominated in marked degree. Unfortunately, only a very few subjects—6—were involved in Flugel's investigation, and, of course, immediate introspections on affective tone were not taken. Duration, furthermore, rather than the number of experiences dealt with as units was the basis of comparison in the study.

A research by Wohlgemuth (31), while inconclusive in its results, points in the same direction as Flugel's. The former's subjects (school children) were requested to recall all they could of the pleasant and unpleasant events of a certain day after intervals of one day and eleven days, respectively. These students recorded that they looked with favor upon most of the happenings. The day, however, of whose events Wohlgemuth requested a chronicle had been for all of his subjects a holiday and, hence, one likely to be more than usually satisfying. Allowing, moreover, as it does, about 12 to 24 hours between the impressions and the first recall of them, the study does not safeguard itself against the criticism that affective influences, if potent, had already had opportunity to operate, when the first records were taken. Wohlgemuth meets this objection, in part, by basing, on a comparison merely of the results of the first-day and eleventh-day recall, his conclusion that there is no consistent difference in the proportion of agreeable and disagreeable experiences remembered.

In those techniques in which the emotional reaction is described for the first time when the experience is recalled one faces another sort of problem—namely, that of insuring that the emotional quality reported is actually that of the original experience rather than of merely the memory of it.

Still another factor which frequently escapes control, especially in those procedures involving the *spontaneous* recall of *remote* events, is that of mood. It is difficult, in other words, to afford each possible memory item the opportunity of standing upon the merits of its original affective components. The temporary mood or set of the recall period may determine the items to be reinstated and mask memories associated with other moods and of greater fundamental survival value.

Those studies in which the subjects learned, or responded in some other way, to lists of words selected by the experimenter as symbolic of pleasant or unpleasant experiences one must also question for several reasons. (*a*) With individual differences enormous, we have little assurance that the words have a like significance for the subjects as for the experimenter; or (*b*) that they stimulated the affective reactions which confessedly they symbolize. (*c*) Little attention is usually given in the construction of the contrasting pleasant and unpleasant lists to the degree of familiarity of the words or to such factors of form as might determine recall value. (*d*) If the speed of associative response, as in a controlled association test, is the criterion of readiness of recall, there is also the danger of slipping into the error of ignoring the fact that associations aroused by unpleasant verbal stimuli may be such as the subject would be embarassed to report and that under these circumstances he would delay until a socially less discreditable response occurred to him.

While avoiding most of the difficulties already mentioned, that procedure which involves the testing of the strength of the associations established between symbols and pleasant or unpleasant odors probably also has limitations in that the associations are rather artificial. It is doubtful whether the symbols presented simultaneously with odors for one time enter into any sort of vital integration with the emotional patterns aroused by the latter. Odors, too, we should expect to be much less provocative of persistent defense reactions than more distinctly social stimuli, a point of which the psychoanalysts might make much.

PROCEDURE

The technique employed in the present study avoids many of the difficulties enumerated in the foregoing discussion. The items, whose retention was to be tested, were intimately bound up with personal and social values, as they were the grades made by students on 10 quizzes in a course in educational psychology.

Ten-minute examinations were given at the end of each of the first 10 class periods in the course, 23 days elapsing between the first and last examination. Each examination paper was carefully graded on the percentage basis and returned to the student at the meeting of the class following the one at which the test was taken. Since the answers in the quizzes were of the yes-no type, few characteristics other than the grade on the paper probably were highly differentiating. The grades, written in red, were kept a constant size as well as always placed in the same corner of the paper. Each student was requested at the end of each class hour to return his graded paper to the instructor and to indicate upon it by means of a rating on a five-point scale the nature of his emotional reaction to his mark. If he was thrilled or very happy over it, he was to rate the paper 1; if utterly disgusted and discouraged, 5; if little concerned one way or the other, 3. Mild reactions in the direction of pleasure and displeasure were represented by ratings of 2 and 4, respectively. The method insures at least a reasonably accurate report of the affective reaction aroused by the stimulus of the grade. Of course, absolute honesty and sincerity were encouraged in the report given concerning the emotional reaction excited. The purpose of the experiment, naturally, was not explained to the subjects until the investigation was completed. There is clear evidence, however, that the uniqueness of our procedures did not prevent genuine concern on the part of the participants over their scholastic standing.

Five weeks after the affective reaction to the last of the 10-minute quizzes had been recorded, a period during which the students were given no numerical grade of any sort and in which no reference was made to examinations, a recall of the 10 grades was requested. All subjects also took at this time the Otis Self-Administering Test of Mental Ability, Form A, and filled out the Woodworth-Wells Personal Data Sheet.

Since no mention had been made of the undesirability of jotting down the marks received, lest such comment prejudice our case, it was necessary to get a statement from each subject in regard to whether or not he had made a memorandum of his ratings, and if so, the length of time since he last referred to this. The subjects were assured that the instructor had a complete record of all of their grades and that nothing was to be gained by falsifying their reports.

More than half of the students, fortunately, had made no written notations. The data from most of those who kept account of their earnings are included in the study only because it was felt that the group who did not might be more than normally homogeneous with respect to some personality traits which are of significance in the recall activities. Laird (17) and Washburn (29) have already recognized the part played by personality in determining the pattern of recall.

In order to control as far as possible in our results the influence of the factors of primacy, recency, and ultimacy, very difficult and very easy examinations were alternated. In one class, Group I, furthermore, the alternation sequence began with an easy examination; in the other, Group II, with a difficult. The two classes of students—highly similar sophomore groups—were given the same lectures and reading assignments as well as the same examinations upon the material covered in the course. An analysis of the distribution of the affective ratings throughout the examination sequence indicates that the alternation scheme employed did apportion any given rating approximately equally among the various tests.

To prevent the repetition of any grade in the 10 which a single subject received, lest this repetition facilitate recall, it was necessary, since the final ratings were always given on a percentage basis, to include a different number of questions in each examination. The percentage method of scoring was employed because it seemed the one that would indicate to the student most unambiguously the degree of his success on an examination.

In the grades given Groups I and II, unfortunately, there were many containing round numbers and five's. Since these grades might be more readily recalled, because of their familiarity, than those involving other number combinations, and since inspection revealed that the former did not occur with exactly the same relative frequency among the papers receiving the different affective ratings, it was necessary to run a check experiment with another group—Group III—highly similar in its composition to Groups I and II. Although graded on a percentage basis just as the other students had been, these received no marks containing a zero or a five. This group began the alternation of difficult and easy quizzes with the former type.

In the ensuing discussion the data gathered from the students in the three classes, I, I, and III, will be presented separately in order to make possible an evaluation of the internal consistency of our findings. A further division, moreover, will be made, namely, that between the returns from the students who did and who did not keep a written record of their grades. The former groups will be given in addition to the class numeral the designation R. The R groups are homogeneous to the extent that none of the students in them referred to his grade record in the five weeks preceding the recall test. Results will be presented for 76 subjects. Of these, only 29 belong in the R groups.

Statement and Analysis of Results

Our data will be analyzed in three ways in order to insure that we are not dealing with the expressions of such factors as primacy, recency, intelligence, or familiarity of number combination. Table 1 presents in summary form the incidence of correctly recalled grades among those receiving the affective ratings of various degrees. Table 2 was formulated to meet the possible criticism that the trends shown in Table 1 may reflect an intelligence rather than a feeling factor. It might be contended, for example, that more pleasing than displeasing grades will be recalled because the high grades, and hence the pleasing ones, will have been earned mostly by the most intelligent students, who, in turn, by virtue of their intelligence, will have recalled more of their grades than the dull students. Although any individual subject's record may be influenced by other than affective variables, the strength of the foregoing contention will be seriously undermined if a much larger number of students recall correctly a greater proportion of their pleasing than displeasing grades.

Table 3, giving the results from the examinations occurring at different stages in the sequence, should enable us to determine whether the same trends are shown throughout. That factors of primacy, recency, etc., might weigh heavily is suggested by the fact that a period of 23 days separated the first and last quiz. The scheme of alternating difficult and easy examinations, to be sure, was used to effect an equal distribution of the given affective ratings through the various stages of the examination sequence; but its success, while marked, was not entire, a fact which might create slight misgivings with respect to the meaning of the data in Table 1.

Although the two more analytical methods of presenting our findings have their limitations—i.e., the individual measures in Table 3 cannot be taken too seriously because of the paucity of cases involved at times, and not all subjects could be used in all the comparisons of Table 2 because every individual did not report affective ratings of every degree—the more consistent patterns appearing in these analyses, especially when considered in relation to those exhibited when the data are massed as they are in Table 1, are not to be treated lightly.

An inspection of the summary table, Table 1, reveals the following relationships: (*a*) A larger percentage of those grades receiving the affective rating 1 are correctly recalled than of those receiving any other rating. The one exception occurs in the case of a group which includes only 5 subjects. (*b*) Among the marks given a rating of 2 by those students who kept no written records of their scholastic earnings the incidence of correctly recalled grades is greater than among those marks given a rating of 3, 4, or 5. In the R groups, however, the percentage of correctly recalled grades tends to be greater for those rated 5 than either 2, 3, or 4, though the differences at times are slight. (*c*) The grades reported as provoking

TABLE 1

TOTAL NUMBERS AND PERCENTAGES OF GRADES OF VARIOUS DEGREES OF PLEASANTNESS AND UNPLEASANTNESS WHICH WERE CORRECTLY RECALLED

Measures	Groups							
	I	II	III	I R	II R	III R	I, II, III	I R, II R, III R
Number of subjects	16	16	15	10	14	5	47	29
Total number of grades rated 5	63	47	69	37	33	22	179	92
Total number of correctly recalled grades among those rated 5	28	18	11	19	15	6	57	40
Percentage of correctly recalled grades among those rated 5	44	38	16	51	45	27	32±.024	43±.035
Total number of grades rated 4	32	27	16	19	19	10	75	48
Total number of correctly recalled grades among those rated 4	10	11	3	9	8	1	24	18
Percentage of correctly recalled grades among those rated 4	31	41	19	47	42	10	32±.037	38±.047
Total number of grades rated 3	20	27	14	11	21	3	61	35
Total number of correctly recalled grades among those rated 3	8	10	1	3	8	1	19	12
Percentage of correctly recalled grades among those rated 3	40	37	7	27	38	33	31±.040	34±.054
Total number of grades rated 2	21	22	22	17	29	5	65	51
Total number of correctly recalled grades among those rated 2	10	12	5	6	10	3	27	19
Percentage of correctly recalled grades among those rated 2	48	55	23	35	34	60	41±.041	37±.046
Total number of grades rated 1	21	40	23	13	35	9	84	57
Total number of correctly recalled grades among those rated 1	13	23	8	10	24	3	44	37
Percentage of correctly recalled grades among those rated 1	62	58	35	77	71	33	52±.037	65±.043

TABLE 2

COMPARISON IN TERMS OF GREATER, LESS, OR SAME, OF THE RELATIVE FREQUENCY OF CORRECTLY RECALLED GRADES AMONG THOSE RECEIVING THE VARIOUS AFFECTIVE RATINGS—RESULTS STATED IN TERMS OF THE NUMBER OF SUBJECTS WHOSE RECORDS AGREE WITH THE CONDITIONS DESCRIBED IN THE COMPARISONS.

Comparisons	Groups					
	I	II	III	I R	II R	III R
Larger percentage of 1's than 5's	8	8	4	5	9	3
Smaller percentage of 1's than 5's	0	4	1	1	1	0
Same percentage of 1's as 5's	1	3	6	1	1	1
No comparisons possible	7	1	4	3	3	1
Larger percentage of 1's than 3's	4	8	0	3	7	3
Smaller percentage of 1's than 3's	0	0	1	1	2	1
Same percentage of 1's as 3's	2	3	5	3	1	0
No comparisons possible	10	5	9	3	4	1
Larger percentage of 5's than 3's	4	6	2	3	5	2
Smaller percentage of 5's than 3's	1	2	1	2	1	1
Same percentage of 5's as 3's	1	3	4	2	4	1
No comparisons possible	10	5	8	3	4	1
Larger percentage of 2's than 4's	6	6	3	3	1	3
Smaller percentage of 2's than 4's	3	2	1	2	4	0
Same percentage of 2's as 4's	3	2	4	2	6	1
No comparisons possible	4	6	7	3	3	1
Larger percentage of 1's than 4's	6	8	3	3	7	2
Smaller percentage of 1's than 4's	0	0	3	1	0	1
Same percentage of 1's as 4's	5	3	3	3	2	1
No comparisons possible	5	5	6	3	5	1
Larger percentage of 1's than 2's	6	6	2	3	8	1
Smaller percentage of 1's than 2's	2	5	2	2	1	2
Same percentage of 1's as 2's	1	1	4	0	2	0
No comparisons possible	7	4	7	5	3	2
Larger percentage of 2's than 5's	5	6	5	2	2	3
Smaller percentage of 2's than 5's	5	4	4	6	4	1
Same percentage of 2's as 5's	0	2	3	0	6	0
No comparisons possible	6	4	3	2	2	1
Larger percentage of 2's than 3's	4	5	2	4	3	1
Smaller percentage of 2's than 3's	1	2	1	1	3	0
Same percentage of 2's as 3's	3	2	3	1	4	2
No comparisons possible	8	7	9	4	4	2
Larger percentage of 5's than 4's	7	4	3	4	2	3
Smaller percentage of 5's than 4's	4	4	3	4	3	1
Same percentage of 5's as 4's	0	4	5	1	6	1
No comparisons possible	5	4	4	1	3	0
Larger percentage of 4's than 3's	3	3	2	2	5	1
Smaller percentage of 4's than 3's	3	3	0	1	2	1
Same percentage of 4's as 3's	4	4	3	3	1	2
No comparisons possible	6	6	10	4	6	1

TABLE 3

Incidence of Correctly Recalled Grades Among Those Receiving the Various Affective Ratings in the Five Examination Groups

Groups	Ratings	Examinations				
		1 and 2	3 and 4	5 and 6	7 and 8	9 and 10
I	5	73	66	60	33	75
	4	50	0	25	16	16
	3	66	66	33	0	33
	2	75	50	50	20	0
	1	100	100	50	0	75
II	5	33	45	64	33	16
	4	40	75	50	0	0
	3	66	50	60	16	0
	2	80	66	100	50	0
	1	82	80	55	25	50
III	5	20	16	6	21	12
	4	20	0	50	25	0
	3	0	20	0	0	0
	2	14	25	25	25	0
	1	0	50	50	0	75
I R	5	86	88	33	0	33
	4	50	50	60	0	40
	3	33	100	100	0	0
	2	60	16	40	0	0
	1	100	100	100	60	50
II R	5	66	62	44	16	25
	4	100	0	100	33	0
	3	0	50	100	33	29
	2	75	57	50	20	0
	1	75	86	50	50	50
III R	5	0	33	0	40	100
	4	0	33	0	50	0
	3	0	0	0	0	0
	2	0	0	0	100	50
	1	0	100	100	100	0

a strong positive or negative affective reaction tend to be recalled to a greater extent than those reported as stimulating neither pleasure nor displeasure. The same general statement might be made for the mild affective reactions, but the relationship is not so consistently evident as in the case of the strong. (*d*) In 9 out of 12 possible comparisons the grades described as very satisfying or dissatisfying are recalled with greater relative frequency than those described as less intensely toned.

That the foregoing trends are not a mere reflection of the variable intelligence is suggested by the material in Table 2, which shows essentially the same patterns as that of Table 1. In all 6 groups, for example, more individuals than otherwise—37 vs. 7—remembered a larger percentage of the grades they rated 1 than 5. With the exception of one group—an R group—the incidence of correctly recalled grades is greater among those rated 2 than 4. Whereas, furthermore, 47 subjects recalled correctly a larger proportion of the grades they reported very pleasant or unpleasant than of those they reported as provoking no affective reaction, only 13 did the reverse. Though the differences are not so pronounced, the same general description would obtain also in the case of the milder affective reactions.

The data of Table 3, like those of 2, are in fundamental agreement with those of Table 1. Twenty-one out of a possible 30 comparisons, each involving a unit of two successive examinations, show a larger percentage of the most pleasing grades to be correctly recalled than of the most displeasing, whereas two more comparisons show the percentages to be the same. The tendency just characterized seems more consistent among the data from the three groups of students who kept records of their grades than among the data from those who did not. In 70% of the comparisons also the incidence of correct recall is greater among those grades which were most satisfying than among those which were neutral; in 13% the reverse obtains. The most disagreeable grades are remembered with relatively greater frequency than the indifferent grades in 60% of the comparisons and with less relative frequency in 30. The tendency for a larger proportion of grades rated 2 to be remembered better than those rated 4 is not so uniform. The recall frequencies, as far as the affective groups 2 and 4 are concerned, are not the same for the students groups who made notations of the marks they received as for those who did not. This we have previously indicated. In 53% of the comparisons the relative incidence of correct recall among grades rated mildly pleasant is greater than for those rated neutral, while in 24% the opposite prevails. Little difference in the extent of recall is revealed by a comparison of the affective groups 3 and 4.

The analyses thus far have shown our data rather consistent internally and, hence, probably not highly colored by such spurious influences as intelligence or temporal position factors. It is worthy of note, furthermore, that the findings for Group 3 are in most respects similar to those for Groups 1 and 2. This is significant, since, as we have already pointed out, the familiarity of the number combinations involved in the grades was not so well controlled as it might have been in the case of the latter groups. In Group 3, however, an attempt was made in the direction of better control of this factor in that zero's and five's were not employed. Of course, it may be argued that even though the material to be remembered by Group III contained no round numbers or five's, still the number combinations in the upper range of the percentile gamut are more familiar than those in the lower, because of certain prevailing sales techniques and because students who survive to the college level must in their academic experience have received more of the higher grades than the lower. Only one paper in Group III, however, was graded the familiar 98, the lower numbers between 90 and 100 occurring much more frequently than the higher. It seems questionable to the author, too, whether the numbers between 90 and 100 are more familiar than those between 0 and 30, for instance. Since, furthermore, for Groups I, II, III, the grades exciting mildly pleasant reactions and hence falling largely outside the range within which the relatively familiar grade or bargain-counter sign combinations occur show essentially the same relationships between recall and affective ratings as do the grades exciting the more vigorous pleasure responses and limited almost entirely to the highest decile, it seems highly probable that our results are not to be interpreted in terms of inequalities in the familiarity of the number combinations receiving the various affective ratings. The fact, moreover, that grades announced as mildly pleasing and displeasing tend to be recalled with slightly greater frequency than those reported as neutral does not fit readily into the familiarity hypothesis.

A correlation type of attack on our data also gives evidence in harmony with the hypothesis of affective influences in recall. The extent of a student's recall of all his grades as well as his recall of unpleasant grades has been studied in relation to his scores on the Woodworth-Wells Personal Data Sheet and on the Otis Self-Administering Test of Mental Ability, Form A. Since the Woodworth-Wells test, it is claimed, can be used as a rough index of the extent of the neurotic tendencies which a subject possesses, and since the vigorous use of defense mechanisms, among which we should expect to find the forgetting of the unpleasant, is alleged to characterize those with neurotic leanings, it seems reasonable to anticipate some correlations of interest in our material.

The correlations given in Table 4 have all been transmuted into

TABLE 4

ZERO-ORDER AND SOME FIRST-ORDER PARTIAL CORRELATIONS BETWEEN THE OTIS, WOODWORTH-WELLS, AND RECALL TESTS—SUBSCRIPT 1 REPRESENTS THE OTIS TEST; 2, THE WOODWORTH-WELLS; 3, THE PERCENTAGE OF RECALL OF ALL GRADES; AND 4, THE PERCENTAGE OF RECALL OF ALL UNPLEASANT GRADES.

	Groups					
Correlations	I	II	III	I R	II R	Av.
r_{23}*	+.17±.17	+.32±.16	+.44±.15	—.49±.17	—.05±.20	+.08
r_{24}*	+.29±.16	+.42±.14	+.27±.17	—.15±.22	+.10±.19	+.19
r_{13}†	+.33±.16	—.36±.15	+.36±.16	+.30±.20	+.23±.18	+.17
r_{14}†	+.04±.18	—.36±.15	—.05±.18	+.36±.20	—.11±.19	—.02
r_{12}‡	+.09±.18	—.13±.17	+.57±.12	+.16±.22	+.28±.17	+.19
$r_{23\cdot1}$	+.15±.17	+.30±.16	+.32±.16	—.59±.15	—.12±.18	+.01
$r_{24\cdot1}$	+.29±.16	+.40±.14	+.36±.16	—.23±.21	+.13±.18	+.19

*A positive correlation means that the highest scores on the Woodworth-Wells Test tend to be associated with the greatest difficulty in recall, and vice versa.

†A positive correlation means that the highest Otis Test scores tend to be associated with the most extensive recall, and vice versa.

‡A positive correlation means that the lowest Woodworth-Wells scores tend to be associated with the highest Otis Test scores, and vice versa.

terms of Spearman's rho. They have been computed separately for each group because there did not exist that entire identicalness in the handling of the groups which would justify massing all measures. Group III R will not be dealt with in the presentation because of its limited membership. Since all of the groups, furthermore, claim under 20 members, and since we have already indicated the danger of considering any one individual's record too seriously because the grades with the different affective tones were not always in the case of a single record distributed equally throughout the examination series, the correlations can at best have only suggestive value. Their uniformities probably have significance.

For the sake of brevity we shall use in the correlation subscript 1 to indicate the Otis test; 2, the Woodworth-Wells test; 3, the percentage of all grades correctly recalled; and 4, the percentage of all unpleasant grades correctly recalled.

The correlations of Table 4 show the following relationships: (*a*) In Groups I, II, and III the correlations between the ranks according to the results of Woodworth-Wells test and the ranks according to the percentages of all grades recalled as well as the percentages of only unpleasant grades recalled are positive, though seldom 3 times the probable error of r in size. In the R groups, the

r's tend to be negative. (*b*) The r_{24}'s are, for the most part, of greater positive value or smaller negative than the r_{23}'s. (*c*) The correlations between ranks according to the Otis test scores and the number of grades remembered are as a rule positive and of moderate size, whereas between the former and the extent of recall of only unpleasant grades they are generally of small positive value or of negative. (*d*) The correlations between the Otis test and the Woodworth-Wells are in the main positive but small; i.e., there is a slight tendency for the brighter pupils to show the fewer traits of possible neurotic significance, and vice versa. (*e*) Holding Factor 1 constant tends to lower by a very small amount the size of the positive correlations between Factors 2 and 3 or to increase the size of the negative, while it tends to alter the correlations between 2 and 4 even less. (*f*) In general, r_{24} is greater or more uniformly positive than r_{14}—i.e., a low rating on the Woodworth-Wells test for the groups we are investigating is more closely associated with effective recall than is a high intelligence test score.

The meaning of the correlations, which are, in the main, small and not very significant, is not entirely obvious. An interpretation in terms of escape adjustments of one sort or another, however, may be given items *b, c, e,* and *f*. Item *d* also fits nicely into such a picture, as the most unintelligent in a university group one would expect to be the most maladjusted, generally speaking—i.e., to have the most to escape. The negative r_{23}'s and r_{24}'s among the R groups, however, suggest, if the Woodworth-Wells score is an index of personal and social maladjustment tendencies, that there may also be a relation between these tendencies and an inclination to dwell on one's failures. In fact, the very recording of grades by a student—the behavior which was the basis for his inclusion in an R group—may be evidence of an apprehensive set on his part. The most apprehensive individuals in a group selected as the R's were, we should expect to magnify failure most or at least to be the most keenly aware of the exact degree of it and vice versa. The correlations of Table 4, then, taken in the large, prompt one to believe that emotional-social imbalance may be associated with extreme trends either in the form of proneness to obliterate memories of failure or to absorption with thoughts of defeat.

An effort to analyze our data further with a view to throwing more light upon the pattern of the forgetting that occurred was unsuccessful because of our inability to devise a sufficiently objective method. We essayed, for instance, to discover whether among those grades falsely reported there was a tendency for the values to be higher than those actually received, i.e., whether unpleasant grades tended to be relegated rather directly to oblivion or to persist transformed in such a way as not to be derogatory to the subject. About all that can be said is that there was certainly no very clearly defined trend in the latter direction. Among the R groups there seemed a slight disposition toward self-underestimation.

Discussion

Although our total findings do not contribute crucially to the problem of the more fundamental mechanisms responsible for the pattern of recall-feeling relationships, it is clear that it would be difficult to account for our results in terms merely of an endopsychic censor, since both pleasant and unpleasant grades seem to be better remembered than those neutrally colored. It is worthy of note, also, that the studies of Smith (24) and of Gordon (8) corroborate this finding. Indeed, it seems questionable, in the light of the variety of trends shown among our data, whether any one principle is adequate for describing, much less explaining them. To expand the point somewhat, among the various views of unpleasantness that have been held—(*a*) that it is a psychic element, (*b*) an expression of the domination of the sympathetic nervous system over the craniosacral, (*c*) a pattern yielded by thalamic activity of a certain sort, (*d*) an instinctive reaction to thwarting, (*e*) an accompaniment of an energy deficiency; (*f*) a withdrawal reaction, (*g*) a value judgment upon stimuli, in terms of whether they normally arouse negatively adaptive responses—the latter seems to the author the most in keeping with the facts, since consistent patterns of body response or uniform mental constellations have not been revealed in the so-called displeasure states or attitudes. If, moreover, concept *g* is correct, we should have little reason to assume that in the specific integration patterns subsumed under the heading of unpleasant activities there would be found any great constancy in those phases which are significant from the point of view of recall. More specifically, some of our subjects may have built up for one reason or another rather generalized habits of reacting to failure or success—i.e., some may review frequently their defeats, some their triumphs. On the other hand, more circumscribed determinants may be operative. A low grade, for example, may give rise to an emotionally more intense response in an individual than a high one because the former symbolizes elimination from college or loss of a Phi Beta Kappa rating, whereas the latter merely makes for a continuation of the *status quo*. For some students, moreover, low grades are unusual experiences and hence vivid; for others the reverse is true. Even friends, since they would be more likely to remind one of his successes than failures, may influence the relative frequency with which pleasant and unpleasant grades are reinstated. The reporting of grades to one's associates, furthermore—a happening which one might more reasonably expect to occur if the marks are creditable than if the reverse were true—may through the nature of the verbalization process fix firmly in mind the items proclaimed. Affective forces, in other words, may be indirectly as well as directly operative. In fact, it is the relatively indirect influences, the author feels, which are responsible for the likeness between her own finding, namely, that unpleasant grades are more readily forgotten than the pleasant—and that which the psychoanalyst announces. In such studies as Gordon's (8), in which social influences were eliminated in large measure, pleasant and unpleasant impressions were remembered in equal degree, but retention was better for both of these classes than for those impressions accompanied by no or very limited affective reactions. Let it be noted that in regard to the relative infrequency of the recall of indifferent experiences, our own results bear out Gordon's.

That what we descriptively label emotional vividness is a prominent factor in recall-affection relationships, then, seems clear; but the general problem of the relative memory value of unpleasant and pleasant experiences can not be settled in terms of so simple a principle. Nor is it to be settled in terms of such hypostatizations as an endopsychic censor, an instinct of self-assertion, an instinct of repression, or a prestige drive. The problem will, doubtless, be given a more satisfactory answer after we have probed deeply into the question of the genesis and development of our values and the relation of these to such extra-organismal influences as social, moral, and philosophical traditions. Even the application to the problem of such depersonalized descriptive concepts as relative congruity, relative frequency, relative vividness, relative extensiveness of associations, etc.—fundamental and invaluable as these are—would leave us minus the first links in the chain and would probably not furnish the most workable formulation of the facts, i.e., the formulation most useful for prediction and control.

SUMMARY

1) Pleasing grades, in general, were correctly recalled relatively more frequently than displeasing ones. Individual subjects, however, seemed to differ in their recall habits in the degree to which they emphasized success and failure.

2) Grades reported as pleasant or unpleasant tended to be better remembered than those reported as stimulating no affective reaction.

3) There seemed to be a slight relationship between the possession of such traits as are listed in the Woodworth-Wells Personal Data Sheet and the tendency to remember relatively poorly unpleasant grades.

4) It is probable that no one principle can account for the findings of the study. The explanation of the trends revealed will very likely be found in both intra- and extra-organismal factors.

REFERENCES

1. ANDERSON, A. C., & BOLTON, F. J. Inhibition of the unpleasant. *J. Abn. & Soc. Psychol.,* 1925-1926, **20**, 300-303.
2. BAXTER, M. F., YAMADA, K., & WASHBURN, M. F. Directed recall of pleasant and unpleasant experiences. *Amer. J. Psychol.,* 1917, **28**, 155-157.
3. CHANEY, R. M., & LAUER, A. R. The influence of affective tone on learning and retention. *J. Educ. Psychol.,* 1929, **20**, 287-291.
4. COLEGROVE, F. W. Memory: an inductive study. New York: Holt, 1920. Pp. 367.
5. FLUGEL, J. C. A quantitive study of feeling and emotion in every day life. Paper read before Brit. Psychol. Soc., March 24, 1917.
6. FOX, C. The influence of subjective preference on memory. *Brit. J. Psychol.,* 1923, **13**, 398-404.
7. GORDON, K. Über das Gedächtnis für affektiv bestimmte Eindrücke. *Arch. f. d. ges. Psychol.,* 1905, **4**, 437-458.
8. ———. The recollection of pleasant and unpleasant odors. *J. Exper. Psychol.,* 1925, **8**, 225-240.
9. ———. A study of early memories. *J. Delinq.,* 1928, **12**, 129-132.
10. GRIFFITTS, C. H. Results of some experiments on affection, distribution of associations, and recall. *J. Exper. Psychol.,* 1920, **3**, 447-464.
11. HENDERSON, E. N. Do we forget the disagreeable? *J. Phil., Psychol., & Sci. Method,* 1911, **8**, 432-438.
12. HOLLINGWORTH, H. L. The oblivescence of the disagreeable *J. Phil., Psychol., & Sci. Method,* 1910, **7**, 709-715.
13. KOHS, S. C. The association method in its relation to the complex and complex indicators. *Amer. J. Psychol.,* 1914, **25**, 544-594.
14. KOWALEWSKI, A. Studien zur Psychologie des Pessimismus. Wiesbaden: Bergmann, 1904. Pp. 122.

15. ————. Arthur Schopenhauer und seine Weltanschauung. Halle a/S, 1908.

16. Külpe, O. Bemerkungen zu vorstehender Abhandlung. *Arch. f. d. ges. Psychol.,* 1905, **4**, 459-464.

17. Laird, D. A. The influence of likes and dislikes on memory as related to personality. *J. Exper. Psychol.,* 1923, **6**, 294-303.

18. Morgan, E., Mull, H. K., & Washburn, M. F. An attempt to test moods or temperaments of cheerfulness and depression by directed recall of emotionally toned experiences. *Amer. J. Psychol.,* 1919, **30**, 302-304.

19. Myers, G. C. Affective factors in recall. *J. Phil., Psychol., & Sci. Method,* 1915, **12**, 85-92.

20. Peters, W. Gefühl und Erinnerung. Beiträge Erinnerungsanalyse. *Psychol. Arbeit.,* 1911, **6**, 197-260.

21. Peters, W., & Nemeck, O. Massenversuche über Erinnerungsassoziationen. *Fortschr. der Psychol. u. Ihrer Anwendungen,* 1914, **2**, 226-245.

22. Peterson, J. The effect of attitude on immediate and delayed reproduction: a class experiment. *J. Educ. Psychol.,* 1916, **7**, 523-532.

23. Ranschburg, P. Über Hemmung gleichzeitiger Reizwirkungen. *Zsch. f. Psychol.,* 1902, **30**, 39-86.

24. Smith, W. W. Experiments on memory and feeling tone. *Brit. J. Psychol.,* 1921, **11**, 236-250.

25. ————. The measurement of emotion. New York: Harcourt, Brace, 1922. Pp. 184.

26. Tait, W. D. The effect of psycho-physical attitudes on memory. *J. Abn. & Soc. Psychol.,* 1913-1914, **8**, 10-37.

27. Tolman, E. C. Retroactive inhibition as affected by conditions of learning. *Psychol. Monog.,* 1917, **25**, No. 107. Pp. 50.

28. Tolman, E. C., & Johnson, I. A note on association time and feeling. *Amer. J. Psychol.,* 1918, **29**, 187-196.

29. Washburn, M. F., Harding, L., Simons, H., & Tomlinson, D. Further experiments on directed recall as a test of cheerful and depressed temperaments. *Amer. J. Psychol.,* 1925, **36**, 454-455.

30. Washburn, M. F., Grang, F., Ives, M., & Pollack, M. Memory revival of emotions as a test of emotional and phlegmatic temperaments. *Amer. J. Psychol.,* 1925, **36**, 456-458.

31. Wohlgemuth, A. The influence of feeling on memory. *Brit. J. Psychol.,* 1922-1923, **13**, 405-417.

32. Wuerdemand, W. Über die Bedeutung des Gefühle für das Behalten und Erinnern. *Neue Psychol. Stud.,* 1926, **1**, 507-573.

University of Texas
Austin, Texas

L'INFLUENCE DE QUELQUES FACTEURS AFFECTIFS SUR LE RAPPEL

(Résumé)

Il a été difficile d'étudier expérimentalement la valeur de rappel des expériences agréables et désagréables parce que les stimuli employés dans le laboratoire ont été insuffisants au point de vue affectif, et dans les situations non du laboratoire il n'a été que rarement possible d'obtenir le rapport du sujet sur la qualité affective de ses expériences pendant leur progrès. Notre étude a probablement conquis ces difficultés en employant comme matière de rappel les notes de quelques examens subis dans un cours de psychologie éducationnelle et en faisant noter leurs sentiments par les étudiants chaque jour quand ils ont reçu leurs examens notés.

Soixante-seize étudiants ont participé à l'expérience. Ils ont subi dix examens, et le rappel des notes reçues à chacun de ceux-ci a eu lieu cinq semaines après le renvoi du dernier examen.

On peut résumer nos résultats ainsi: (1) Les notes agréables ont été correctement rappelées généralement plus fréquemment que les désagréables; et les notes de ton affectif que celles de ton neutre. (2) Les mémoires des sujets individuels ont semblé différer selon le degré d'importance du succès ou de l'insuccès. (3) Il semble qu'il y ait eu un petit rapport entre la possession des traits donnés dans la Feuille de Traits Personnels Woodworth-Wells et la tendance à se rappeler avec une difficulté relativement grande les notes relativement désagréables.

Il est probable qu'aucun principe seul ne peut rendre compte de nos résultats. Il est certain que le concepte du censeur endo-psychique ne peut le faire.

Koch

DER EINFLUSS EINIGER AFFEKTIVER FAKTOREN AUF ERINNERUNG

(Referat)

Es ist bisher sehr schwer gewesen, den Erinnerungswert von lust- und unlustbetonten Erfahrungen experimentell zu untersuchen, weil die im Laboratorium verfügbaren Reize affektiv unzureichend waren, und in andern Situationen ausserhalb des Laboratoriums ist es selten möglich gewesen, den Bericht einer Person über die affektive Eigenschaft ihrer sich im Gang befindlichen Erfahrungen zu erhalten. Unsere Untersuchung hat diese Schwierigkeit vermutlich überwunden, indem als das zu erinnernde Material die Zensuren von Klausuren genommen wurden, die im Kurse von Erziehungspsychologie gegeben wurden, und indem man die Studenten ihre Gefühle bei der Rückgabe der schriftlichen Arbeiten schätzen liess.

76 Studenten nahmen am Experiment teil. Es wurden zehn Klausuren gegeben und fünf Wochen nach der Rückgabe der letzten Arbeit wurde nach den Zensuren aller Arbeiten gefragt.

Unsere Ergebnisse lassen nich folgendermassen zusummenfassen: (1) Angenehme Zensuren wurden im Allgemeinen verhältnismässig häufiger richtig erinnert als unangenehme; und affektiv betonte Zensuren häufiger als unbetonte.

(2) Die individuellen Erinnerungen schienen sich in dem Grade zu unterscheiden, als Erfolg oder Misserfolg betont waren.

(3) Es schien eine geringfügige Beziehung zwischen solchen Zügen zu bestehen, wie sie auf den Woodworth-Wells Blättern für personliche Data aufgeführt sind, und der Tendenz verhältnismässig schlechte unangenehme Zensuren zu erinnern.

Es erweist sich als wahrscheinlich, dass nicht ausschliesslich ein Prinzip unsere Ergebnisse erklären kann. Gewiss hingegen ist, dass die Auffassung vom endo-psychischen Zensoren der Aufgabe nicht gerecht wird.

KOCH

AN EMPIRICAL STUDY OF THREE PSYCHOPHYSICAL METHODS*[1]

From the Psychological Laboratories of the University of Chicago

Kate Hevner

The Problem

It was the object of this study to compare, on the basis of empirical data, some of the commonly used methods of constructing psychological scales. In order to make such a comparison, it would be necessary to construct the psychological scales from the same material, with the same subjects, and under conditions similar in every respect except the method by which the material was presented to the subject for judging. By such a procedure we can isolate any phenomena which are not inherent in the nature of the material, or in the criteria on which the individual judgments are made, but which are characteristic only of the method of making the judgments. We can learn whether the scale values resulting from the use of the various methods are comparable, and, if they are not, we may be able, by further analysis, to learn which series of values most nearly approximates the true values.

The three methods selected for study are methods which originated in the early psychophysical experiments on lengths of lines, intensities of gray values, lifted weights, etc., where the material could be measured both objectively and subjectively. These methods were later applied to material which could be measured by subjective methods only, and for which no adequate objective measure could be found.

Many such applications of the older psychophysical methods have been made by psychologists who are interested in converting a qualitative series into a series of quantitative differences. The work was begun in America by Cattell, who used the order of merit method for his ratings of scientific men (1). His methods and variations

---

*Recommended by L. L. Thurstone and received for publication by Carl Murchison of the Editorial Board, April 29, 1929.

[1] The writer is indebted to Professor L. L. Thurstone of the University of Chicago for the suggestion of this problem and for supervision in carrying out the study.

of them were subsequently applied to other fields of subjective measurements. The method of equal appearing intervals was used by Thorndike (2) in order to calculate quantitative scale values for a series of subjective differences. Thorndike's methods have also been applied to other fields. These older methods differ from the methods used in this study in the number of items which are judged, in the number of individual judgments made for each item, and in the details of procedure, but more especially in the statistical treatment of the data. The law of comparative judgment, introduced by Thurstone in 1927 (3), has been used in the calculation of these scale values, rather than the older psychophysical processes which involve the use of the probable error as a unit of measurement. The value of this process in calculating scale values, for very complex stimuli which involve a great deal of personal prejudice and bias, has already been demonstrated (3). Its validity for calculating the scale values for the handwriting specimens used in this experiment is further justified by the internal consistency of these data. The average discrepancy between the calculated proportions of each judgment, and the corresponding experimental proportions, which serves as a measure of internal consistency, is less than .03.

The three methods chosen for study were the order of merit method, the method of paired comparison, and the method of equal appearing intervals. The process applied to the first two methods was that of the law of comparative judgment in its simplest form, *Case V*. A graphical process was used for calculating the scale values in the method of equal appearing intervals. All the details of these procedures are explained in the section on "The Experimental Procedure."

The material chosen for scaling was handwriting. A comparison could therefore be made of the methods when used on psychological values instead of on simple physical stimuli. Since a study of method only was intended, it might be argued that more simple and homogenous material should have been used, e.g., some material which would involve only one variable, such as length of line, intensity of gray value, extent of areas, etc. But the advantage of simplicity, with material of this sort, is offset by the disadvantages of presentation. Such material must be presented by the experimenter in the laboratory to each subject individually, and the amount of time involved would be prohibitive. Since the method was of primary interest, it was essential that there should be a large number of judgments on each stimulus, so that the stability of the results should be unquestionable. It seemed necessary, therefore, to prepare material which could be manipulated by the subjects outside the laboratory, and without the help or supervision of the experimenter. Handwriting specimens satisfied these conditions very well, and carried with them the added advantage that they could not be actually weighed or measured by unreliable subjects.

To the layman, the criteria for judging excellence in handwriting are almost entirely subjective. Personal preferences and abilities are involved, as well as the legibility and the pleasing appearance of the writing. Such complexity is typical of most of the variables to which these methods of measuring are being applied at the present time, for example, English composition, specimens of sewing, and other handiwork, nationality and race preferences, and attitudes of various kinds. That the distribution of opinion and preference in regard to these variables follows the normal curve, must be proved in each individual case by the internal consistency of the data, and in this case, with excellence in handwriting as the variable, the proof seems clear.

After some preliminary experimentation, it was decided to define the handwriting variable more closely by asking the subjects to base their judgments on three criteria—neatness, uniformity of the slant, and uniformity of the stems and ovals of the letters. These three criteria were suggested twice in the instructions to the subjects, and there was also a warning to avoid such factors as the "character" which the writing showed, and the particular personal associations which might be called out by certain specimens.

The Experimental Procedure

Three hundred and twenty individuals were asked to write the phrase "A very good explanation" in their ordinary handwriting on the penciled line of a 4 x 6 card. Those individuals were chosen whose handwriting would give as wide a range as possible in degree of excellence. They included university students, high school and grade school pupils, unskilled laborers, bookkeepers, and handwriting experts. Ten adult subjects then sorted these 320 samples into seven piles, placing all the worst specimens on the pile at the extreme right, all the best specimens on the pile at the extreme left, all the average specimens on the pile in the middle, and arranging the other four piles so that all seven should form a graded series of excellence in handwriting. The instructions to the judges in this preliminary experiment were merely: "Judge each specimen by whatever makes it seem good or bad handwriting to you." The judges were allowed to sort and resort as much as they wished, until they were entirely satisfied with their judgments. They consumed on the average an hour and a half in completing their work for this preliminary experiment.

From the results of these experiments, the average position among the piles was reckoned for each of the 320 specimens, and the specimens arranged in rank order on the basis of these averages. Then, beginning with the best specimens and working through the whole series to the worst specimens, approximately every fourth card was selected in order to secure the final series of 72 specimens which was used in the method of equal appearing intervals. Since it was planned to make a scale of *slant* writing only, the specimens which inclined more to the vertical were avoided. Certain specimens in which the quality of the ink was not well adapted to the zincographing process were also discarded. From this series of 72, again approximately every fourth specimen was selected in order to secure another graded series of 20 specimens for the order of merit method. Code numbers were assigned to the 72 specimens, and they were reproduced by a zincographing process on slips of white paper, 1½ x 4 inches. The zincographing process reduced the original size of the handwriting very slightly, but otherwise provided a faithful reproduction.

The 20 specimens used in the order of merit method were also used in the method of paired comparison. The code numbers were cut off, and the 190 necessary pairs of specimens were arranged in two parallel columns, on 8½ x 11-inch pages. Each pair of specimens was set off by heavy black lines in order to reduce the error from side comparisons. The 12 pages containing these specimens were stitched together in the upper left hand corner, and a rotating order was used with them so that the effect of fatigue would not always come on the same pages.

All the material for the experiment was included in a heavy manila envelope, 9 x 12 inches, which could be placed in the subject's hands without further instruction or supervision from the experimenter. The material included (*a*) the 12 pages of specimens for the method of paired comparison; (*b*) an 8½ x 4-inch envelope containing the paper clips, the covering slips labelled *a, b, c, d* . . . .

k, and the 72 specimens for the order of merit method, and (*c*) the printed page of instructions. The instructions read as follows:

This envelope contains a number of specimens of adult slant writing from which a scale of handwriting is to be constructed. The scale will be made by three different psychophysical methods. In each of the methods you will be asked to compare the different samples and judge the quality of the writing. By "quality" we mean "good appearance," that is, neatness, uniformity of the slant, and uniformity of the stems and ovals of the letters. Base your judgments on these factors, and be careful to eliminate such factors as the "character" which the writing shows, or the particular associations, pleasant or unpleasant, which you may have with certain specimens. The scale is to be made on the basis of:

1) Neatness.
2) Uniformity of the slant.
3) Uniformity of the stems and ovals of the letters.

Method 1

1) You are given eleven slips with letters on them, *A, B, C, D, E, F, G, H, I, J, K*. Please arrange these before you in regular order. On slip *A* put those specimens which you judge to be the *best* specimens. On slip *F* put those specimens which you judge to be the *average* specimens. On slip *K* put those specimens which you judge to be the *worst* specimens. On the rest of the slips, arrange the specimens in accordance with the degree of excellence you find in them.

2) This means that when you are through sorting you will have eleven piles arranged in order from *A*, the best, to *K*, the worst.

3) Do not try to get the same number on each pile. They are not evenly distributed.

4) The numbers on the slips are code numbers, and have nothing to do with the sorting.

5) You will find it easier to sort them if you look over a number of specimens chosen at random before you begin to sort.

6) It will probably take you about 20 minutes to sort them.

7) When you are through sorting, please clip the piles together, each with its letter slip on top. Replace the 11 piles, clipped carefully, in the envelope marked "Method 1," and seal.

Method 2

In this method, spread out the 20 samples before you and arrange them in sequence according to their merit. Select the best specimen and place it at the extreme left; then find the second best specimen and place it beside the first, and the third best and place it beside the second, and so on down to the worst specimen, which will appear at the extreme right. When the whole sequence is laid out before you, you may wish to make some revisions. When you have finished, gather the specimens together in one pile, with the best one on top and the worst one on the bottom. Place the rubber band around the whole pile and seal them in envelope marked "Method 2."

Method 3

Use the 12 pages of specimens and follow directions given at the top of the first page.

The subject arranged the specimens according to the instructions, fastened the piles securely with paper clips and rubber bands, and sealed them in the proper envelopes. Approximately one hour was spent by each subject in performing this part of the experiment.

The experiment was performed by graduate and undergraduate students in the University of Chicago, and by undergraduate students in Wilson College, Chambersburg, Pennsylvania, in Colgate University, Hamilton, New York, and in the State College for Teachers, Superior, Wisconsin. Omissions and mistakes made it necessary to eliminate the work of some of these subjects, but the experiment was completely and correctly performed by 370 subjects.[2]

The Statistical Treatment of the Results

The results for the method of equal appearing intervals are given in Table 1. This table shows the number of times each specimen appeared in each of the 11 piles. Records are included for each of the 72 specimens, although in this study scale values have been calculated only for those specimens which are also included in the order of merit method and the paired comparison method. In Table 1, these 20 specimens are marked with an asterisk.

The results for the method of paired comparison could be tabulated directly from the 12 pages of specimens, and these results are recorded in Table 2.

Table 3 shows a similar set of data for the order of merit method, although this tabulation could not be made directly from the subject's records, as in the case of the paired comparison method. The rank order of the specimens for each subject was first recorded, and especially prepared data sheets were then used to tabulate for each subject the number of times each specimen was preferred to every

---

[2] The writer wishes to acknowledge her indebtedness to all the students who gave their time as subjects for the experiment, and more especially to their instructors for the generous cooperation in distributing and collecting materials—to Professor Louise Holcomb and Miss Vera Oberschelp of Wilson College, to Professor Paul L. Whitely of Colgate University, to Miss Isabel King Wallace of the Superior State College, and to Mr. W. L. Sharp of the University of Chicago.

FREQUENCY DISTRIBUTIONS FOR EACH OF THE 72 SPECIMENS
METHOD OF EQUAL APPEARING INTERVALS

	1	2	3	4	5	6	7	8	9	10	11		1	2	3	4	5	6	7	8	9	10	11
a2			1	1	5	15	50	90	104	81	23	*n3*		1		3	6	22	54	86	95	77	26
b2	6	14	34	39	65	96	56	43	16		1	*p3*							20	19	73	116	142
c2				2	2	7	26	67	95	129	42	*q3*								1	9	29	331
**d2*								1	6	40	323	*r3*		6	16	31	42	74	90	45	47	17	2
e2	2	9	16	28	51	73	73	59	37	15	7	*s3*	118	81	72	37	28	24	7	2	1		
**f2*	6	18	29	54	65	91	64	30	8	5		**t3*	2	1	2	3	8	20	51	94	95	73	21
g2	2	10	25	48	56	93	58	51	21	6		**u3*				1			3	7	44	144	171
h2	3	15	25	35	40	83	66	53	38	14		*v3*					2	1	1	7	42	145	172
i2	70	109	65	48	37	28	8	2	1	1	1	*w3*	46	76	78	60	51	34	15	8	2		
**j2*				5	8	27	57	85	100	78	10	**x3*			1	3	3	19	35	50	110	119	30
**k2*	118	83	53	46	26	31	7	2	3	1		*y3*	270	62	21	7	3	7					
m2	12	53	68	70	59	52	29	18	5	3	1	**z3*			5	7	13	49	61	82	86	54	13
**n2*	2	7	18	27	38	76	91	46	44	18	3	**a4*					2		4	11	42	145	166
p2	1	11	34	52	56	97	46	40	17	13	3	**b4*	2	9	23	30	41	79	71	61	37	14	3
q2							4	16	32	135	193	*c4*	1	1	2	9	18	60	77	78	79	40	5
**r2*			4	1	5	12	14	51	79	134	70	*d4*			1	1			1	4	16	93	254
s2	65	76	60	60	52	30	18	5	3	1		*e4*	3	6	20	33	50	86	87	46	29	9	1
t2	1	9	18	34	57	94	78	39	28	7	5	**f4*	80	87	68	50	34	38	7	3	1	2	
u2	7	6	31	56	63	72	74	42	17	8	1	**g4*			1			7	3	11	39	96	213
v2	51	83	79	55	41	41	9	8	1	1	1	*h4*	1	1	6	22	31	62	79	71	66	29	2
w2	2	5	27	36	57	79	70	51	28	12	3	*j4*	1		3	2	7	27	56	75	82	89	28
x2		5	8	17	28	75	76	79	55	33	4	*k4*	8	38	35	58	63	85	38	26	14	4	1
y2			4	4	9	34	50	83	90	82	14	*m4*	2		9	13	18	62	82	92	55	33	4
z2		3	14	28	47	89	81	54	37	14	3	*n4*									3	23	344
a3										8	362	*p4*	79	88	87	39	22	34	7	12	1	1	
b3	23	56	58	70	55	64	29	7	6	2		*q4*		6	9	18	28	70	90	63	62	22	2
c3	10	24	39	54	47	95	41	37	13	10		*r4*	52	73	72	52	46	55	10	6	3	1	
d3	11	26	42	63	59	71	51	28	14	4	1	*s4*	6	27	43	59	57	87	46	28	12	5	
e3	4	27	51	48	47	103	58	21	8	3		**t4*	79	89	57	54	34	27	12	11	6	1	
f3	2	4	13	18	33	61	73	74	63	22	7	*u4*	34	84	91	58	50	30	11	5	3	1	
**g3*	95	83	59	41	40	39	7	2	3	1		*w4*	3	15	46	67	55	88	62	23	10	1	
**h3*	38	82	82	59	40	50	10	4	5			*x4*	7	7	22	25	42	72	60	67	50	15	3
i3	26	68	81	54	63	46	20	8	2	2		*z4*	65	76	82	66	28	35	8	4	3	3	
**j3*						1		3	8	65	293	*a5*	1	2	5	9	16	56	77	76	77	44	7
k3	126	102	57	26	20	31	5	1	2			**b5*	157	76	54	32	20	22	7	2			
m3				2	9	24	61	92	97	67	18	**y5*	73	62	59	49	43	46	18	11	7	2	

TABLE 2

The Number of Times the Specimen at the Top was Preferred to the Specimen at the Side in a Total of 370 Judgments
Method of Paired Comparison

	b5	*y5*	*t4*	*f4*	*k2*	*g3*	*h3*	*n2*	*f2*	*j2*	*z3*	*b4*	*x3*	*r2*	*t3*	*g4*	*u3*	*a4*	*j3*	*d2*
b5		124	55	121	118	120	148	29	41	4	5	9	1	6	12	2	1	0	0	0
y5	246		135	148	199	208	151	33	38	9	17	22	4	2	13	6	1	1	0	0
t4	315	235		259	237	281	294	122	205	24	23	106	26	11	56	4	5	0	0	0
f4	249	222	111		189	205	215	62	46	2	11	14	3	8	12	5	0	0	0	0
k2	252	171	133	181		191	207	70	92	9	17	41	7	6	13	4	3	2	0	0
g3	250	162	89	165	179		157	43	64	1	7	31	0	4	1	2	0	0	0	0
h3	222	219	76	155	163	213		50	38	2	12	10	2	6	18	3	1	0	0	0
n2	341	337	248	308	300	327	320		177	31	32	143	20	16	39	4	2	2	0	1
f2	329	332	165	324	278	306	332	193		18	46	99	9	7	25	5	2	1	0	0
j2	366	361	346	368	361	369	368	339	352		146	303	89	69	226	36	24	1	1	2
z3	365	353	347	359	353	363	358	338	324	224		278	67	40	192	16	13	23	0	2
b4	361	348	264	356	329	339	360	227	271	67	92		31	25	122	10	11	2	0	0
x3	369	366	344	367	363	370	368	350	361	281	303	339		170	321	63	54	10	0	2
r2	364	368	359	362	364	366	364	354	363	301	330	345	200		295	29	73	67	7	3
t3	358	357	314	358	357	369	352	331	345	144	178	248	49	75		25	8	11	3	3
g4	368	364	366	365	366	368	367	366	365	334	354	360	307	341	345		191	224	58	58
u3	369	369	365	370	367	370	369	368	368	346	357	359	316	297	362	179		153	33	12
a4	370	369	370	370	368	370	370	368	369	369	347	368	360	303	359	146	217		16	21
j3	370	370	370	370	370	370	370	370	370	369	370	370	370	363	367	312	337	334		145
d2	370	370	370	370	370	370	370	369	370	368	368	370	368	367	367	312	358	349	225	

TABLE 3

The Number of Times the Specimen at the Top was Preferred to the Specimen at the Side in a Total of 370 Judgments
Order of Merit Method

	b5	*y5*	*t4*	*f4*	*k2*	*g3*	*h3*	*n2*	*f2*	*j2*	*z3*	*b4*	*x3*	*r2*	*t3*	*g4*	*u3*	*a4*	*j3*	*d2*
b5		95	97	107	138	141	101	9	30	1	5	14	0	0	3	0	0	0	0	0
y5	275		194	202	216	220	196	44	71	9	14	38	7	5	8	2	1	0	0	0
t4	273	176		179	212	214	204	44	86	8	9	26	4	0	10	0	0	0	0	0
f4	263	168	191		206	210	185	25	41	3	5	15	3	2	4	1	1	0	0	0
k2	232	154	158	164		165	147	26	49	2	6	15	0	1	3	0	1	1	0	0
g3	229	150	156	160	205		154	21	43	1	3	19	1	3	3	1	1	0	0	0
h3	269	174	166	185	223	216		29	55	3	3	12	2	0	5	1	0	1	0	0
n2	361	326	326	345	344	349	341		275	60	84	166	25	15	35	1	0	3	0	0
f2	340	299	284	329	321	327	315	95		14	28	81	7	8	9	2	1	0	0	0
j2	369	361	362	367	368	369	367	310	356		213	329	104	72	183	14	11	5	0	0
z3	365	356	361	365	364	367	367	286	342	157		282	73	38	146	7	5	3	1	1
b4	356	332	344	355	355	351	358	204	289	41	88		34	15	70	2	3	3	0	0
x3	370	363	366	367	370	369	368	345	363	266	297	336		131	252	18	20	21	1	2
r2	370	365	370	368	369	367	370	355	362	298	332	355	239		296	27	50	63	15	5
t3	367	362	360	366	367	367	365	335	361	187	224	300	118	74		7	8	9	2	1
g4	370	368	370	369	370	369	369	369	368	356	363	368	352	343	363		249	233	105	75
u3	370	369	370	369	369	369	370	370	369	359	365	367	350	320	362	121		176	55	31
a4	370	370	370	370	369	370	369	367	370	365	367	367	349	307	361	137	194		25	31
j3	370	370	370	370	370	370	370	370	370	370	369	370	369	355	368	265	315	345		141
d2	370	370	370	370	370	370	370	370	370	370	369	370	368	365	369	295	339	339	229	

TABLE 4

Scale Values for the 20 Specimens from All Three Methods

Specimen	Scale values Equal appearing intervals	Order of merit	Paired comparison
*b*5	1.6	.0000	.0000
*g*3	2.15	.4548	.4431
*h*3	2.8	.6471	.4510
*y*5	2.8	.8593	.4664
*f*4	2.25	.6160	.5548
*k*2	1.8	.4336	.6679
*t*4	2.35	.8041	1.4315
*f*2	5.13	1.9442	1.7219
*n*2	6.23	2.6589	1.8829
*b*4	6.1	2.8801	2.4367
*t*3	8.1	4.2595	3.2037
*z*3	7.7	3.7929	3.3046
*j*2	8.1	4.2171	3.5206
*x*3	8.7	4.9039	4.3280
*r*2	9.3	5.4426	4.4369
*u*3	9.95	6.8529	5.5545
*a*4	9.9	6.7741	5.5302
*g*4	10.15	7.3983	5.8731
*j*3	10.4	8.2750	7.4342
*d*2	10.5	8.5833	7.6887

other specimen. The results of this method, in their final form are, therefore, comparable to the results of the method of paired comparison, and the statistical treatment of the results was the same for both methods.

The next step in the study was the calculation of a scale value for each specimen, in each method. In the method of equal appearing intervals the calculation of the scale value was a comparatively simple procedure. From the frequency distribution for each specimen, a cumulative frequency curve was plotted, and the median, the point at which the curve crossed the 50% line, was read directly from the graph. The medians for each of the 20 specimens are given in Table 4.

Since the psychophysical law of comparative judgment can be applied to any situation in which a number of observers make one discriminatory judgment for each possible pair of stimuli in any given stimulus series, this law was applied to the data from both the paired comparison and the order of merit methods. It would be assumed under this law that the distribution of preference in handwriting is normal on the psychological continuum, and this assumption may be questioned. However, there is already some evidence that the assumption of a normal distribution is justified in the case of a very complex continuum such as nationality preferences (4), as well as for the simple sense distances of the more formal psychophysical experiments; and the assumption, in the case of handwriting excellence, is undoubtedly a safe one, since there is an average discrepancy of less than .03, when the calculated proportion of each judgment is compared with the corresponding experimental proportion.

The law in complete form is as follows (4, p. 407):

$$S_1-S_2=X_{12}\sqrt{\sigma_1^2+\sigma_2^2-2r_{12}\cdot\sigma_1\cdot\sigma_2} \qquad (1)$$

in which

S_1-S_2=the scale distance between the two modal discriminal processes

X_{12}=the sigma value corresponding to the observed proportion of judgments, "Stimulus$_1$ is greater than Stimulus$_2$." The proportion is designated $p_{1>2}$. The numerical value of X_{12} is positive when $p_{1>2}$ is greater than .50, and it is negative when $p_{1>2}$ is less than .50.

σ_1 and σ_2=ambiguities (discriminal errors) of the two stimuli.

r_{12}=the correlation between the discriminal deviations involved in the same judgment.

In this study we have used *Case V* in which the law is applied to a group of subjects, and each subject gives only one judgment for every stimulus pair. In *Case V* it is assumed that $r_{12}=0$, and that the value of σ is unity for every stimulus in the series. With these approximations, the law takes the simpler form:

$$S_1-S_2=X_{12}\sqrt{2}, \qquad (2)$$

when σ is chosen as the unit of measurement.

From Table 2 which gives the data for the method of paired comparison, a corresponding table of proportions was made, and from this table, the percentage of judges who preferred each specimen to every other specimen could be read directly. The tentative rank order for each specimen could also be easily calculated from this table by summing the percentages for each of the specimens in turn.

The next step was to prepare a table of sigma values corresponding to the percentages. Since a procedure of weighting all the proportions in such a calculation as this would be extremely laborious, the most unreliable proportions were not used in calculating the scale values. Those which were dropped were proportions of .97 or above and of .03 or below. Those proportions which were retained were all given equal weight. When the measures of internal consistency were calculated, it was apparent that these approximations were not unwarranted.

From this table of sigma values the scale separation for every consecutive pair of specimens was calculated. Theoretically it would be possible to calculate the scale separation for any pair of specimens in the table, no matter how far apart they might be on the scale, but actually such a procedure would not be practical. When two specimens are widely separated on the scale, many of the proportions are included between zero and .03 and between .97 and 1.00, and are dropped out of the calculations. The scale separation would be calculated from the few remaining pairs of sigma values, and therefore could not be considered satisfactory. For this reason, the scale separations have been calculated only for those specimens which were adjacent in rank order.

It was necessary to prepare a separate table for the calculation of each of the scale separations. The data for these tables was obtained from the table of sigma values, and with these 19 tables the procedure in general was as follows:

Let us suppose that the two specimens for which we are to calculate the separation are the 11th and 12th specimens when arranged in rank order. They will be designated then in general terms by S_{11k} and S_{12k}, respectively. If any of the other specimens be designated by S_k, then from equation (2) page 201 we have

$$S_{11k} - S_k = X_{11k}\sqrt{2} \qquad (3)$$

and

$$S_{12k} - S_k = X_{12k}\sqrt{2} \qquad (4)$$

where $X_{11k}\sqrt{2}$ represents the distance on the scale between specimen S_{11k} and the other given specimen S_k, and $X_{12k}\sqrt{2}$ represents the distance on the scale between S_{12k} and the given specimen.

Subtracting (4) from (3) we have

$$S_{11k} - S_{12k} = \sqrt{2}(X_{11k} - X_{12k}). \qquad (5)$$

In other words, $\sqrt{2}(X_{11k} - X_{12k})$ is the scale separation of S_{11k} and S_{12k} when calculated in terms of but one of the 20 specimens.

Summing such values for all the specimens, we have, in general terms,

$$n(S_{11k}-S_{12k})=\sqrt{2}\Sigma[X_{11k}-X_{12k}] \qquad (6)$$

$$\text{or} \quad (S_{11k}-S_{12k})=\frac{\sqrt{2}}{n}[\Sigma X_{11k}-\Sigma X_{12k}] \qquad (7)$$

that is, the actual scale separation for any two specimens is equal to the sum of the scale separations for those two in terms of all the other specimens, divided by the total number of such calculated scale separations.

By means of these 19 calculated scale separations the handwriting scale was constructed. Specimen *b*5, which was preferred more often than any other specimen was arbitrarily chosen as the origin for the handwriting scale, and the scale values of all the other specimens were procured by successive additions of the scale separations. Since all the other specimens were preferred less often than *b*5, and are therefore lower in a scale of handwriting excellence, a numerically high scale value indicates poor handwriting. The scale values for the 20 specimens by the method of paired comparison are given in Table 4, and they run from .00 to 7.69. The better specimens have the lower numerical values.

After the scale values have been ascertained, it is necessary to apply a test for the internal consistency of the data. This can be done by using the scale values of the specimens as starting points, and by reversing the whole process, and finally arriving at a set of calculated proportions which may be compared with the experimental proportions. Since we now know the scale values, we can calculate the value of the X_{12} in the equation

$$S_1-S_2=X_{12}\sqrt{2},$$

and with the use of the probability table, this value may be converted into the corresponding proportion. The degree of similarity between such calculated proportions and the actual experimental proportions will give a measure of the internal consistency of the data.

Twenty tables (one for each specimen) were made in order to calculate this consistency. The average discrepancy between the calculated and the experimental proportions ranged from .038 to .00015, and the average discrepancy for the whole table was .0239.

The procedure for calculating the scale separations and final scale values for the specimens in the order of merit method was exactly the same in every respect as the procedure in the case of the method

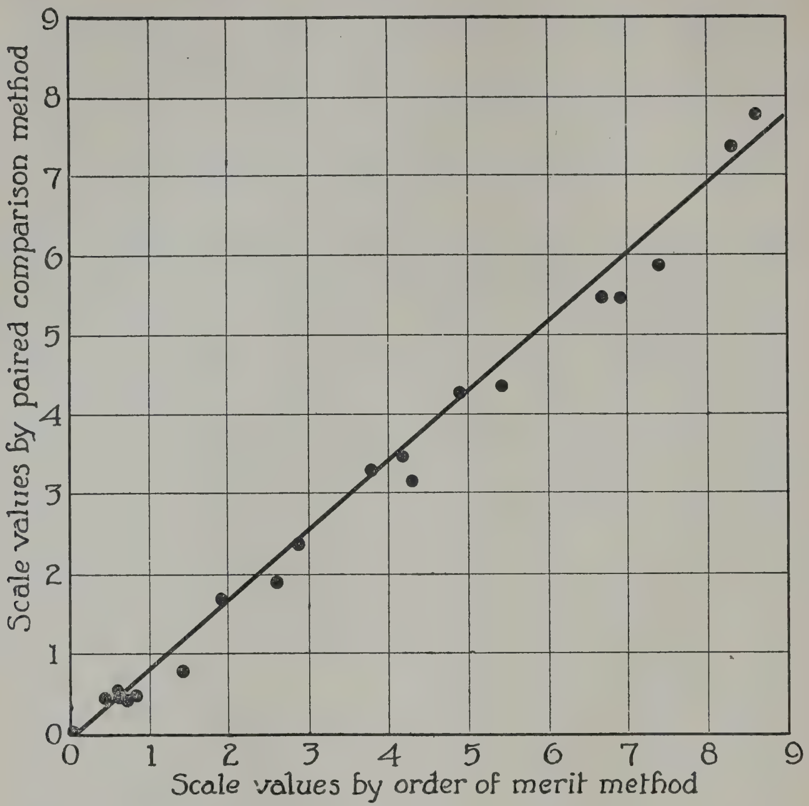


FIGURE 1
Comparison of Scale Values from the Order of Merit and the Paired Comparison Methods

of paired comparison. The scale values are given in Table 4, and the average discrepancy between the calculated and the experimental proportions was .012.

Comparison of the Results

When the scale values for all the specimens have been calculated for each method, the methods can be compared by plotting the scale values for one method against the values obtained from each of the other two. This has been done in Figures 1, 2, and 3. The data from which these figures have been made may be found in

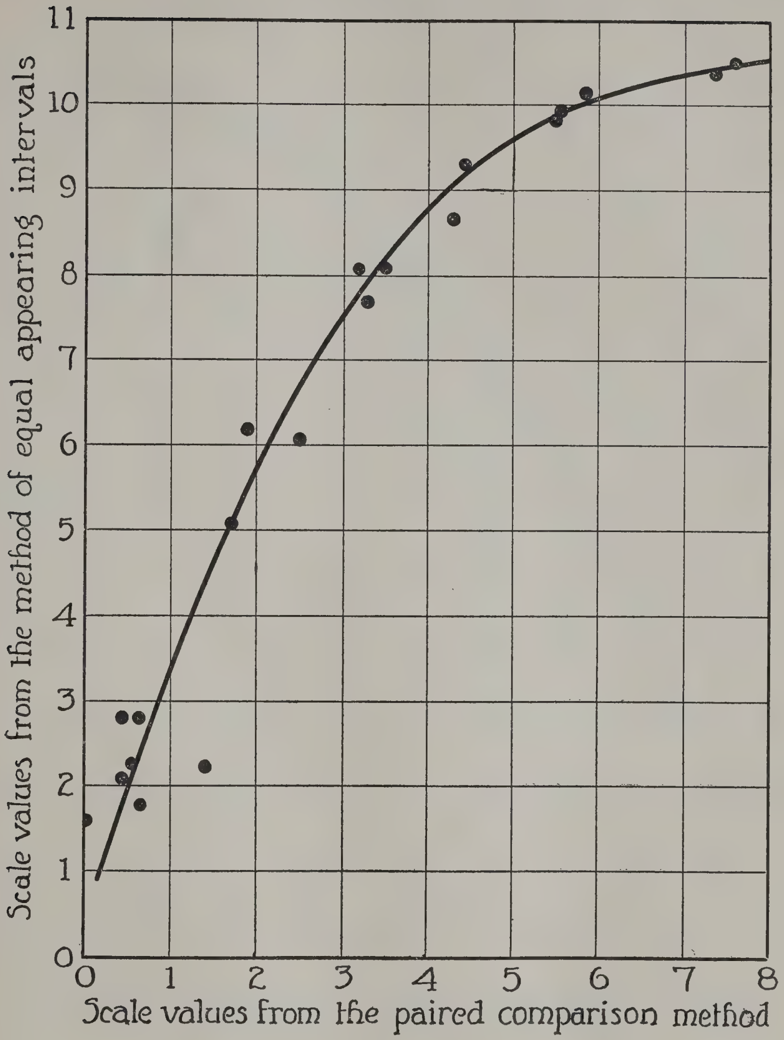


FIGURE 2

Comparison of Scale Values from the Method of Equal Appearing Intervals and the Paired Comparison Method

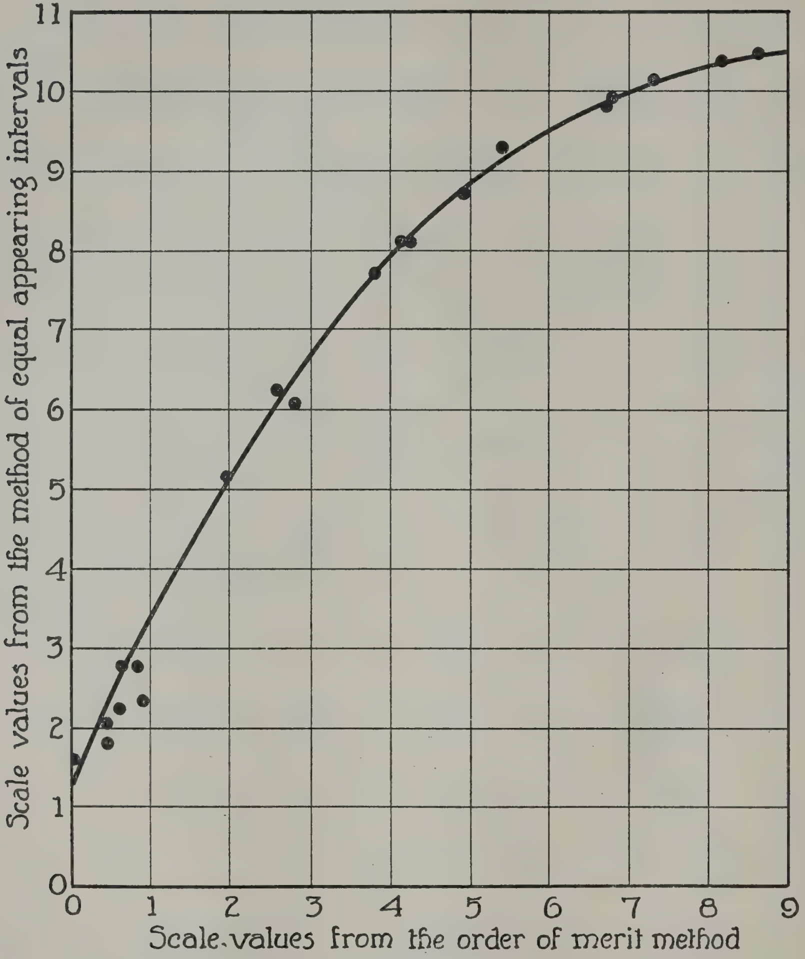


FIGURE 3

Comparison of Scale Values from the Method of Equal Appearing Intervals and the Order of Merit Method

Table 4. In Figure 1, which represents a comparison of the method of paired comparison and the order of merit method, it is apparent that the relation between the two sets of scale values is a linear one. If it be assumed that the correlation between these two sets of data is unity, their regression line will take the form of

$$y=\frac{y}{x}-\frac{y}{x}M_x-M_y$$

and substituting the actual values of the means and the sigmas, the equation for the line in Figure 1 will be

$$y=1.18x-.009$$

An inspection shows that this line fits these data very well.

From the practical point of view it is valuable to know that the two methods yield the same results. The "set up" for an experiment with the method of paired comparison is much more elaborate and costly, the method takes more of the subject's time, and with certain kinds of material is almost impossible to apply. The order of merit method is much simpler, requires less material and less of the subject's time, but the collection of the data in order to apply the law of comparative judgment is a very laborious process for the experimenter.

An inspection of Figures 2 and 3 shows that the relation between the method of equal appearing intervals and each of the other two methods is not linear. In each of these figures the best fitting line has been approximated by inspection, and the relation is clearly curvilinear. The similarity of these two curves confirms the results illustrated in Figure 1. The handwriting scale derived from either the order of merit method or the method of paired comparison would be practically the same scale, but a scale derived from the method of equal appearing intervals would be quite different.

In considering the question as to which of these scales is the most reliable and valid, there are several factors which point to the superiority of the method of paired comparison and the order of merit method over the method of equal appearing intervals.

1) The method of paired comparison and the order of merit method give approximately the same scale values, that is, the relationship between them is linear.

2) The scale values from the method of equal appearing intervals are different from those of either of the other two methods. The relationship in each case is curvilinear.

3) In the method of paired comparison and the order of merit method the internal consistency of the data has been clearly demonstrated. In the method of equal appearing intervals no test for such consistency has been applied and the absence of such a test must, for the present, be counted as the most important argument against it.

4) In the method of equal appearing intervals the frequency distributions for those specimens at the extremes of the scale are not normal distributions (see Table 1). They suffer from an "end effect" which produces a skewed curve. This effect cannot be avoided since it is in the nature of the method that the scale must be arbitrarily cut off at either end, but it may be possible to devise a statistical treatment which would correct for it. Then the corrected medians for these skewed distributions would be as representative of the quality of the handwriting at the extremes of the scale as are the medians for those specimens near the center of the scale, where the range is not arbitrarily restricted.

5) The "just noticeable difference" or the discriminal error which is the fundamental psychophysical unit of measurement is incorporated in both the order of merit method and the method of paired comparison but it is not involved in the calculation of the scale values in the method of equal appearing intervals. This constitutes a very important difference, and it is surprising in view of it that the methods agree as well as they do.

It should be possible to work out a psychophysical process which would take into account the discriminal error for each specimen in calculating its scale value in the method of equal appearing intervals as well as in the other two. The invention of such a process, together with a test for internal consistency is undoubtedly the next step in the analysis of this problem. The assumption under *Case V* of the law of comparative judgment, that the sigma values for all the specimens are equal, is another point of departure for further analysis, and until these studies are completed the final decision on the validity and reliability of these methods cannot be pronounced.

SUMMARY

The object of this study was to compare the scale values of 20 specimens of handwriting, from data obtained by the paired comparison method, the order of merit method, and the method of equal appearing intervals. The judgments of handwriting excellence were obtained for all three methods from the same 370 subjects.

The scale values for the method of equal appearing intervals were calculated by plotting the cumulative frequency curve from the data for each specimen and reading the value of the median directly from the graph. These scale values are given in Table 4.

The scale values in the method of paired comparison and in the order of merit method were obtained by applying the law of comparative judgment in its simplest form, *Case V*. These scale values are given in Table 4. The internal consistency of these calculations is shown by the very small discrepancies between the experimental and the calculated proportions. The average discrepancy is .0239 for the paired comparison method and .012 for the order of merit method.

The results show that the scale of excellence in handwriting would be the same whether the data were obtained by the method of paired comparison or by the order of merit method. The relation between the scale values as calculated by these two methods is linear (see Figure 1).

The relationship between the scale values calculated by the method of equal appearing intervals and either of these other two methods is curvilinear and the handwriting scale constructed by the method of equal appearing intervals differs from the scale constructed by either of the other two methods.

In their present state of development and from the point of view of validity and reliability, the scales constructed by either the order of merit method or the method of paired comparison seem to be superior to the scale constructed by the method of equal appearing intervals, because of the following facts:

1) The relationship between the scale values from the two former methods is linear.

2) The relationship between the values from the method of equal appearing intervals and the values from either of the other two methods is curvilinear.

3) The internal consistency has not been demonstrated for the method of equal appearing intervals.

4) There is a source of error in this method from the "end effect."

5) The psychophysical unit, the j.n.d., is not involved in the calculations for the scale values in this method.

As a method of collecting data from which psychological scales are to be constructed, the method of equal appearing intervals cannot at present be evaluated. A process must first be invented which will take account of the j.n.d. in calculating scale values so that these values may be comparable to the values from the other two methods, and a test for internal consistency must be devised.

REFERENCES

1. HOLLINGWORTH, H. L. Professor Cattell's studies by the method of relative position. *Columbia Univ. Contrib. Phil & Psychol.*, **22**, No. 30.
2. THORNDIKE, E. L. Handwriting. *Teach. Coll. Rec.*, 1910, **11**, 1-81.
3. THURSTONE, L. L. A law of comparative judgment. *Psychol. Rev.*, 1927, **34**, 273-286.
4. ————. An experimental study of nationality preferences. *J. Gen. Psychol.*, 1928, **1**, 405-425.

University of Minnesota
Minneapolis, Minnesota

UNE ÉTUDE EXPÉRIMENTALE DE TROIS MÉTHODES PSYCHOPHYSIQUES

(Résumé)

On a fait cette étude dans le but de comparer les valeurs d'échelle de 20 spécimens d'écriture selon des données obtenues par la Méthode de Comparaisons Mises en Paires, la Méthode d'Ordre de Mérite, et la Méthode D'Intervalles Egaux. On a obtenu les évaluations de l'excellence de l'écriture des mêmes 370 sujets pour toutes les trois méthodes.

On a calculé les valeurs d'échelle pour la Méthode d'Intervalles Egaux en faisant la courbe de fréquence comulative selon les données de chaque spécimen, et en lisant la valeur du médian directement du graphique. On donne ces valeurs d'échelle dans la Table 4.

On a obtenu les valeurs d'échelle dans la Méthode de Comparaisons Mises en Paires et la Méthode d'Ordre de Mérite en appliquant la Loi de Jugement Comparatif dans sa plus simple forme, *Cas V*. On donne ces valeurs d'échelle dans la Table 4. La cohérence interne de ces calculations est montrée par les très petites différences entre les proportions expérimentales et calculées. La différence moyenne est de 0,0239 pour la Méthode de Comparaisons Mises en Paires et de 0,012 pour la Méthode d'Ordre de Mérite.

Les résultats montrent que l'échelle d'écriture serait la même si l'on obtenait les données par la Méthode de Comparaisons Mises en Paires ou par la Méthode d'Ordre de Mérite. La relation entre les valeurs d'échelle comme calculées par ces deux méthodes est linéaire (Voir Figure I).

La relation entre les valeurs d'échelle calculées par la Méthode d'Intervalles Egaux et l'une ou l'autre de ces deux méthodes est curviligne, et l'échelle d'écriture construite par la Méthode d'Intervalles Egaux diffère de l'échelle construite par l'une ou l'autre des deux autres méthodes.

Dans l'état actuel de leur développement et au point de vue de la validité et de la constance, les échelles construites par la Méthode d'Ordre de Mérite ou la Méthode de Comparaisons Mises en Paires semblent supérieures à l'échelle construite par la Méthode d'Intervalles Egaux, à cause des faits suivants:

1) La relation entre les valeurs d'échelle des deux premières méthodes est linéaire.

2) La relation entre les valeurs de la méthode d'Intervalles Egaux et les valeurs de l'une ou l'autre des deux autres méthodes est curviligne.

3) La cohérence interne n'a pas été montrée pour la Méthode d'Intervalles Egaux.
4) L' "effet de bout" ("end effect") nous donne une source d'erreurs.
5) L'unité psychophysique, la j. n. d., n'entre pas dans les calculations pour les valeurs d'échelle dans cette méthode.

Comme méthode pour trouver les données selon lesquelles on peut construire des échelles psychologiques, on ne peut actuellement évaluer la Méthode d'Intervalles Egaux. D'abord il faut inventer un procédé qui rendra compte de la j. n. d. en calculant les valeurs d'échelle, de sorte que l'on puisse comparer ces valeurs avec les valeurs des deux autres méthodes, et il faut trouver un test de cohérence interne.

HEVNER

EINE EXPERIMENTELLE UNTERSUCHUNG DREIER PSYCHOPHYSISCHER METHODEN

(Referat)

Die Absicht dieser Untersuchung bestand darin, die Skalenwerte von 20 Handschriftproben zu vergleichen, an Hand von data, die man nach den folgenden Methoden erhielt: die Methode gepaarter Vergleichung (Paired Comparison Method), Methode der Wertordnung (Order of Merit Method) und Methode gleicherscheinender Intervalle (Method of Equal Appearing Intervals). Die Urteile über die Güte der Handschriften erhielt man für alle drei Methoden von denselben 370 Versuchspersonen.

Die Skalenwerte für die Methode gleicherscheinender Intervalle wurden dadurch gewonnen, dass man eine Kurve der kumulierenden Häufigkeit von den Ergebnissen jeder Probe entwarf und den Mittelwert direkt von der graphischen Darstellung ablas. Diese Skalenwerte sind in Tafel 4 gegeben.

Die Skalenwerte der Methoden gepaarter Vergleichung und der Wertordnung wurden dadurch gewonnen, dass das Gesetz vergleichenden Urteils in seiner einfachsten Form angewandt wurde, Fall V. Diese Skalenwerte sind in Tabelle 4 wiedergegeben. Die innere Übereinstimmung dieser Berechnungen ist durch die sehr kleinen Abweichungen zwischen den experimentellen und den berechneten Proportionen erwiesen. Die durchschnittliche Abweichung beträgt 0,0239 für die Methode gepaarter Vergleichung und 0,012 für die der Wertordnung.

Die Ergebnisse zeigen, dass die Wertskala der Handschriften immer dieselbe ist, ob sie durch die Methoden der gepaarten Vergleichung, oder der Wertordnung gewonnen wurden. Die Beziehung zwischen den nach den beiden Methoden berechneten Skalenwerten ist linear. (siehe Fig. 1).

Die Beziehung zwischen den Skalenwerten, wie sie einerseits durch die Methode der gleicherscheinenden Intervalle, und anderseits durch eine der andern zwei Methoden berechnet wurden, ist krummlinig; und die nach der Methode der gleicherscheinenden Intervalle berechnete Handschriftskala ist von der nach irgend einer der zwei andern Methoden konstruierten verschieden.

Nach dem gegenwärtigen Stand der Entwicklung und vom Standpunkt der Gültigkeit und Zuverlässlichkeit, scheinen die nach den Methoden der Wertordnung und der gepaarten Vergleichung berechneten Skalen der nach der Methode gleicherscheinender Intervalle überlegen zu sein, un zwar auf Grund folgender Tatschen: 1) Die Beziehung zwischen den Skalenwerten der ersteren Methoden ist linear. 2) Die Beziehung zwischen den Werten der Methode gleichersheinender Intervalle und den Werten irgend einer der andern beiden ist krummlinig. 3) Die innere Übereinstimmung ist für die Methode gleicherscheinenden Werte nicht erwiesen. 4) Es liegt darin eine Fehlerquelle, die vom "Endeffekt" herrührt. 5) Die Psychophysiche Einheit, die j. n. d. (just noticeable differences=eben bemerkbare Unterschiede) ist bei dieser Methode in den Berechnungen der Wertskalen nicht einbezogen.

Man kann zur Zeit den Wert der Methode der gleicherscheinenden Intervalle, für das Sammeln psychologischer Skalen noch nicht bemessen. Es muss zuerst ein Verfahren erfunden werden, welches die j. n. d. im Berechnen der Skalenwerte miteinbezieht, so dass diese Werte mit den Werten der beiden andern Methoden verglichen werden können, ferner muss ein Test für die innere Übereinstimmung ausgearbeitet werden.

HEVNER

THE INFLUENCE OF LARGE AMOUNTS OF MANUAL GUIDANCE UPON HUMAN MAZE LEARNING*[1]

From the University of Arkansas

R. H. WATERS

Conclusions as to the influence of guidance upon maze learning have, without exception, been based on studies in which the guidance has been confined to the first 20 trials or less. With but one exception, when the influence of initial guidance upon all criteria is considered, the general conclusion is that the larger amounts of guidance are less efficacious than the smaller. The single exception is represented in the investigation of Alonzo, who finds that, with animals, the effectiveness of guidance increases in proportion to the amount given. The study reported in this paper was undertaken to investigate the influence of initial guidance, given manually, when administered in amounts much larger than any so far reported.

Four groups of 20 subjects each were employed. These subjects had had no previous experience in maze learning. The majority of them were sophomores, drawn from classes in general psychology. The few freshmen and juniors were equally distributed throughout the groups. One of these groups, C, constituted the control group and learned the maze in the usual trial-and-error manner. The other three groups were the experimental or guided groups and were given different amounts of manual guidance. These groups will be designated by the symbols *G*-20, *G*-40, *G*-80; the figures referring to the number of initially guided trials given the respective groups.

The guidance was given by grasping the stylus below the subject's hand and guiding it through the correct pathway. Each trip through the maze took the same amount of time for all groups, viz., 20 seconds. An interval of one minute was inserted after the first 20 guided trials, and intervals of 20 seconds after each succeeding

---

*Received for publication by Carl Murchison of the Editorial Board, February 6, 1930.

[1] Research Paper No. 168, Journal Series, University of Arkansas. The writer wishes to acknowledge his indebtedness to Dr. J. A. McGeoch in the preparation of this paper.

10 trials, in order to avoid fatigue. Immediately upon the completion of the guidance the subjects were asked to complete the learning without further assistance.

The maze employed was the "medium maze" previously used and figured by McGeoch and Melton and hence need not be reproduced here. In all cases the subjects were required to complete the mastery of the maze at a single sitting. The subjects who were unable to do so within an hour and a half were considered failures and their records discarded.

The following instructions were given orally to the subjects before they began work.

> "This is an experiment in maze learning. A maze is essentially a board with a number of grooves or alleys cut into it. There is one place in the maze called the beginning place and another, the ending place or goal. There is one connected series of alleys running from the beginning to the goal, this is called the true path. Other alleys lead off from this true pathway and are called blind alleys. You will take this stylus, hold it perpendicularly, and find the true pathway. I will place the stylus at the beginning place and will tell you when you reach the goal. This constitutes one trial. I will place you at the beginning again and you will again trace the true pathway, and so on. When you are able to trace the true pathway four times in succession without an error, i.e., without any entrance into a blind alley and with no retracing, you are through."

For those receiving guidance the following additional instructions were given.

> "I will take hold of the stylus below your hand and will lead you through the maze, over the true pathway, for 20 (or for whatever number of trials was to be given) trials. That is, we will start at the beginning, go through to the goal, and then immediately go to the beginning for a second trial, and so forth. You will allow your arm and hand to be as passive as possible. However, you are expected to pay close attention while the guidance is being given and learn all that you can about the maze, so that, when the guidance is over, you will be able to learn the maze as quickly as possible."

The stylus was guided down the center of the pathway, making as few contacts with the sides of the alley as possible.

Any retracing over any section of the true pathway and any entrance into a section of a cul-de-sac, during the free learning period, were recorded as errors. Time was taken for each trial by means of a stop-watch. Thus the data include trial, cul-de-sac, retracing, and time records.

The comparative data are summarized in Table 1. The records of the guided groups for the period subsequent to guidance only are shown. The first column of this table shows the groups, the remaining columns, the mean number of trials, errors, and time (in seconds), with their respective *S.D.*'s, required to master the maze, and, in the last column, the average final speed based on the average time per trial for the four criterial trials.

TABLE 1

TRIAL, ERROR, AND TIME RECORDS FOR THE CONTROL AND GUIDED GROUPS

	Trials		Errors		Time		Av. final
Group	Mean	*S.D.*	Mean	*S.D.*	Mean	*S.D.*	speed
C	37.48	22.15	466.24	242.24	2191.00	1032.54	28.75
G-20	42.75	23.98	314.45	165.11	1714.55	982.07	28.16
G-40	27.70	14.21	235.90	185.51	1433.75	759.36	36.44
G-80	23.40	9.10	192.75	154.85	1232.60	577.62	38.33

TABLE 2

PERCENTAGES OF GAIN FOR THE GUIDED GROUPS

Group	Trials	Errors	Time	Av. final speed
G-20	—14	33	22	2
G-40	26	49	35	—27
G-80	38	59	44	—33

The percentages of gain for the various experimental groups as compared with the control group are shown in Table 2.

The following conclusions are based on the data shown in Tables 1 and 2: (*a*) The efficacy of the guidance increases with the amount given, as measured by the criteria of trials, errors, and time. (*b*) The efficacy of the guidance does not increase in direct proportion to the amount, i.e., while 40 guided trials are more beneficial than 20, they do not produce double the benefit of 20, with the exception of the trial records. Likewise 80 guided trials do not produce double the benefit of 40. (*c*) The efficacy of the guidance varies with the criterion, affecting the error records most, then the time, then the trial records, and, with the exception of G-20, is detrimental with regard to the average final speed. (*d*) If we consider the average final speed as an indication of the relative degree of mastery of the maze (it at least reflects the subject's assurance of his ability to run the maze), guidance, with the exception of G-20, is detrimental and the deleterious effect increases with the amount given.

Whether or not the reduction in trials, on the part of the guided groups, is significant, can only be determined by noting the average time and average errors per trial for the various groups. For it is entirely possible that, although the average number of trials is decreased, the average time and error records per trial are increased. If such be the case, the significance of the trial reduction, to a large extent, vanishes. This further analysis of the data is presented in Table 3. From this table it is clear that the value of the trial reduction on the part of the guided groups is not marred by a corresponding increase in the average time and error records per trial. In fact, the average time and errors per trial are actually decreased.

In order to discover whether or not guidance was equally beneficial in eliminating both types of errors, the total error records may be analyzed into the cul-de-sac and retracing errors. These data are shown in Table 4. The first column shows the various groups, the remaining ones, the mean number of culs-de-sac and of retracings, with their respective *S.D.*'s, and also the percentage of gain for the guided groups.

TABLE 3

AVERAGE TIME AND AVERAGE ERROR PER TRIAL FOR THE PERIOD SUBSEQUENT TO GUIDANCE

Group	Average error per trial	Percentage gain	Average time per trial	Percentage gain
C	12.44		58.46	
G-20	7.36	41	40.11	31
G-40	8.52	32	51.76	11
G-80	8.24	34	52.68	10

TABLE 4

CUL-DE-SAC AND RETRACING ERRORS FOR THE CONTROL AND GUIDED GROUPS

Group	Culs-de-sac Mean	*S.D.*	Percentage gain	Retracings Mean	*S.D.*	Percentage gain
C	264.95	148.64		201.29	116.04	
G-20	198.35	109.45	25	116.10	66.40	42
G-40	130.90	109.18	51	105.00	80.95	48
G-80	110.80	99.38	58	81.95	59.03	59

These data warrant the conclusions: (*a*) Both types of errors, when separately considered, manifest the same general tendency as the total error records. (*b*) With the exception of group G-20, both culs-de-sac and retracings are equally affected, being diminished, roughly, 50% in the case of the guided groups.

The significance of the tendencies manifested in the above table is enhanced by two further considerations. The first concerns the number of subjects affected. It is conceivable, of course, that a few individuals with extremely good scores might lower the mean for the group, although the rest of the individuals in the group remain relatively unaffected by guidance. Whether or not this is so can be determined by counting the number of subjects in the guided groups who have better scores than the mean of the control group. When this is done, it is found that in every case the majority of the guided subjects show improvement, indicating that guidance is a factor which tends to influence all subjects. The poorer trial record of Group G-20 is due to the poor records of less than half the cases in that group.

The second consideration involves the statistical reliability of the results as shown in Table 5. The first column of this table shows the various groups, the remaining ones, the mean, $S.D._{av.}$, $S.D._{diff.}$,

TABLE 5

MEASURES OF THE RELIABILITY OF THE DIFFERENCES BETWEEN THE GUIDED AND CONTROL GROUPS

Group	Mean	$S.D._{av.}$	$S.D._{diff.}$	*Diff.*	$\frac{Diff.}{S.D._{diff.}}$
		Criterion trials			
C	37.48	4.83			
G-20	42.75	5.36	7.21	5.27	.73
G-40	27.70	3.17	5.78	9.78	1.69
G-80	23.40	2.03	5.24	14.08	2.68
		Total errors			
C	466.24	52.89			
G-20	314.45	36.94	64.50	151.79	2.35
G-40	235.90	41.50	67.23	230.34	3.43
G-80	192.75	34.64	63.22	273.49	4.33
		Time			
C	2191.00	225.44			
G-20	1714.55	219.70	314.78	476.45	1.51
G-40	1433.75	169.88	282.28	757.25	2.50
G-80	1232.60	129.22	259.84	958.40	3.69

the difference, *Diff.*, between the respective guided groups and the control, and the value of the fraction, *Diff.* divided by $S.D._{diff.}$, respectively.

Most of the comparisons in this table show inadequate reliability. However, it should be noted that, with increasing guidance, the reliability of the differences increases consistently until, with the largest amount of guidance, adequate reliability is obtained. This consistent increase indicates that guidance tends to produce a beneficial effect, although for the smaller amounts the effect is small and unreliable.

So far, our comparisons have been based on the records of the groups subsequent to guidance. It has been customary in the past to count guidance as part of the total learning records. In this experiment we measured the amount of guidance given in two ways: (*a*) in terms of trials, and (*b*) in terms of time. This latter was made possible, as was stated earlier, by giving the guided trials at the uniform rate of 20 seconds per trial. Thus the guidance for Group G-20, for example, amounted to 20 trials and 400 seconds. Now, obviously, when we include in our learning records the trials and time spent in guidance, we will get quite different results from those already quoted, as is shown in Table 6. In this table the average trial and time records when guidance is included are shown, together with the percentage of gain or loss for the guided groups. For the sake of comparison, the average trial and time records subsequent to guidance, with the percentage of gain or loss for the experimental groups, are again presented. The error records are not shown since they are not affected by this comparison.

Conclusions as to the influence of large amounts of guidance now

TABLE 6

TOTAL GUIDED AND UNGUIDED TRIAL AND TIME RECORDS

Group	Total trials	Percentage gain	Total time	Percentage gain	Unguided trials	Percentage gain	Unguided time	Percentage gain
C	37.48		2191.00		37.48		2191.00	
G-20	62.75	— 67	2114.55	3	42.75	—14	1714.55	22
G-40	67.70	— 81	2233.75	— 2	27.70	26	1433.75	35
G-80	103.40	—176	2832.60	—29	23.40	38	1232.60	44

read as follows, based on the total guided and unguided records: (*a*) Guidance is detrimental in terms of trials and the deleterious effect increases with the amount. (*b*) In terms of time, only the largest amount of guidance given produces a significant detrimental effect, the two smaller amounts having little effect. (The conclusion as to the effect on the error records remains unaltered, viz., that guidance is beneficial and increases in proportion to the amount.)

While such a comparison as is shown in Table 6 is valid, it should not be overlooked that in making it we are not considering guidance as an independent, antecedent factor, whose influence upon subsequent learning we are seeking to determine, but rather we are thinking of it as being all of a piece with learning. This distinction has not, in earlier papers on guidance, received the attention its importance merits. There are several reasons which lead to this failure: (*a*) The guidance is measured in the same units as is the subsequent learning, and, simply because the two are measured in the same units, we are led to look upon them as being the same thing. (*b*) The activity employed on the part of the subject while being guided is the same, to external appearances, as that which he utilizes in his undirected learning. (*c*) The two, guidance and learning, follow each other without any intermission. In other words, were the guidance measured in units different from those employed in measuring the rate of learning, or were the activity involved during the guidance of an entirely different type from that employed in the subsequent learning, then the guidance would have been looked upon as a prior condition and not considered only as a part of the learning. In short, we must evaluate the efficacy of guidance in two ways: First, as a prior condition; secondly, as to its effect in increasing the total learning records. Thus, when asked as to the influence of guidance, we may say that spending some time in guidance will enable the subject to master the problem in less time and with less effort, but that his total learning time will be increased to such and such an amount. A somewhat analogous situation is found in connection with the distribution of effort in learning. Distributing the effort brings mastery more rapidly when evaluated in terms of the number of repetitions of the material necessary to learn, but the total amount of time spent from the beginning of the problem to its final mastery is increased. Hence, in any particular instance of distribution, the individual must choose which he would rather sacrifice, repetitions or total time. Likewise, in the case under consideration, he must choose whether he will spend so much time in guidance and then have the advantage of more rapid learning at the expense of total time and trials.

It should be noted also that in most of the papers on guidance the time element has been neglected and the guidance has been measured in terms of trials only. Such a method of measurement is of doubtful validity, since the subject is also putting in time and hence the guidance must be measured both in terms of trials and time. In other words, the guidance usually employed in maze learning is a two-dimensional factor and should be so measured.[2]

It is on the basis of such considerations as these that we are emphasizing the influence of guidance as a prior conditioning factor in subsequent learning. However, it is also incumbent upon one to show the resultant effect when the guidance is included as a part of the total learning records, as we have done in Table 6.

Similar results are obtained from a study of the learning curves, which are not included in this paper. In terms of both errors and time, the curves for the experimental groups start at a considerable lower level and, with some fluctuation from the 5th to the 12th trial, remain below that of the control group until about the 27th trial, from which point on the curves are nearly identical for all groups.

A further word might be added with reference to the relative variability of the several groups. Coefficients of variability were computed according to the Pearson formula. They are shown in Table 7. Here again we meet the question relative to the inclusion of guidance in the means for the experimental groups, and again it is necessary to show the relative variability when this is included as well as when subsequent learning records, only, are considered. Hence both sets of values are shown.

With respect to this factor, Table 7 shows (*a*) that the relative variability, in terms of trials and time, decreases with increasing guidance, by both methods of computation, and (*b*) that variability in terms of errors increases with increasing amounts of guidance. It will also be noted that the decrease in variability in terms of trials

---

[2] The question might be raised as to whether or not guidance, in this latter case, is being counted twice, or given double weight, in the total learning records. Possibly this may account for its being measured in one dimension only in earlier papers.

TABLE 7

COEFFICIENTS OF VARIABILITY

Group	Trials subsequent to guidance	Total trials	Total errors	Time subsequent to guidance	Total time
C	59.10	59.10	51.96	47.13	47.13
G-20	56.10	38.22	52.51	57.28	46.44
G-40	51.30	20.99	78.64	52.96	33.99
G-80	38.89	8.80	80.34	46.86	20.39

TABLE 8

MEAN RETENTION AND SAVING SCORES FOR THE CONTROL AND GUIDED GROUPS

Group	Trials		Total errors		Time	
	Mean	Percentage saving	Mean	Percentage saving	Mean	Percentage saving
C	4.33	90	17.33	96	157.95	93
G-20	8.80	81 (87)	20.00	94	267.55	85 (88)
G-40	5.30	83 (93)	9.95	96	218.15	86 (91)
G-80	6.25	77 (94)	34.25	82	258.25	81 (91)

and time is markedly greater when guidance is included in the learning records, as would be expected.

RETENTION

Each of the subjects was asked to return after an interval of seven days for retention tests. In most of the cases they were told that they would be asked to relearn the maze, and in all cases they were asked to refrain from any form of review during the interval. We have no reason to believe that these instructions were not followed. Granted that it is not a highly satisfactory method of controlling the factor of review, still we believe it to be better than giving them no warning.

Each subject, on his return, was told that he would be given the stylus and that his problem consisted in bringing his learning up to the point at which he had left it, namely, four trials in succession without error. No guidance was given to any of the subjects. Similar records were kept as in the original learning and saving scores computed. These are presented in Table 8. Here once more, since guidance may be considered as part of the total learning score, saving scores must be computed in two ways; first, subsequent to guidance, and, secondly, when guidance is included. Consequently both sets of values are given. The table shows the mean retention scores in trials, errors, and time, and the saving scores for the various groups. The figures in parenthesis represent the saving scores when the guided trials and time are included in the learning records.

Considering first the records subsequent to guidance it is seen that: (*a*) By all criteria the control group manifests a higher score than any of the guided groups, save the error record of G-40 which is equal to the control group. (*b*) The smallest saving scores are obtained by the group given the largest amount of guidance, viz., G-80. (*c*) The saving scores for G-20 and G-40 fall midway between the control and G-80 and are approximately equal, the saving of G-40 being very slightly greater than that of G-20.

When the comparison is made with guidance included as a part of the original learning records, however, the saving scores for the criteria of trials and time manifest a different tendency. The saving scores tend to increase with the amount of guidance given, the saving scores of G-40 and G-80 surpassing that of the control group in the trial criterion.

The retention data may also be evaluated in terms of recall scores. This may be done in two ways, (*a*) in terms of the absolute amount retained, based on the average time and errors for the first relearning trial; and (*b*) in terms of the relation between the average time for the first relearning trial and the average time per criterial trial in the original learning. This latter method is justified on the grounds that we may consider the average time per criterial trial as a measure of the final degree of mastery of the various groups and the relation between this and the average time on the first relearning trial as an indication of the relative amounts of loss of the maze habit. These data are presented in Table 9. The first column shows the groups; the second and third the average errors and time, respectively, for the first relearning trial; the fourth column shows the average time per criterial trial in learning; the fifth and sixth show the absolute loss in time and the percentage loss, respectively.

The following conclusions are based on the data presented in this table: (*a*) In absolute amount retained, as measured by the number of errors made on the first relearning trial, the guided groups surpass the control. No consistent relationship, however, exists be-

TABLE 9

Recall Scores for the Control and Guided Groups

Group	Recall scores errors	Recall scores time	Av. final time in learning	Absolute loss	Percentage loss
C	5.62	46.47	28.75	17.72	62
G-20	3.40	37.65	28.16	9.49	34
G-40	2.15	41.35	36.44	4.91	13
G-80	2.44	55.81	38.33	17.48	46

tween amount of guidance and amount of error reduction. (*b*) The absolute amount retained, as measured in terms of the average time per first relearning trial, decreases with the amount of guidance given. But only the group given the most guidance, G-80, exhibits poorer retention than the control. In measuring retention in terms of the relationship between time on the first relearning trial and the final speed in learning, better records are shown by the guided groups, that is, the guided groups lose relatively less than does the control. Again, however, there is no consistent relationship between the amount of loss and the amount of guidance.

The question may be raised as to the bearing of these results on the relationship between economical methods in learning and degree of retention. It is a somewhat general assumption that the best method for learning is also best for retention. This relationship, however, is a function of a number of different factors, as shown by the present experiment. (*a*) Some of these factors enter into the determination of the "best" method for learning. Thus we have shown that, if guidance is considered as a prior condition, its influence upon subsequent learning is beneficial in terms of all three criteria. However, if we consider the average final speed as a measure or indication of the degree of mastery, guidance produces less efficiency than free learning. Again, it guidance is included as part of the total learning, its efficacy is a function of the criterion, being decidedly detrimental in terms of trials, slightly so in time, and favorable in terms of errors. (*b*) A second group of factors deal with the measurement of the degree of retention. It is well known, of course, that the degree of retention varies with the way it is measured. In the present experiment, retention, measured in terms of saving scores, is, in general, higher in the case of the control group. In terms of absolute recall, five of the six possible comparisons favor the guided groups. In terms of the relation between final speed in learning and time on the first relearning trial, results favor the guided groups. Thus the relationship between economical learning and degree of retention is a complex one and its final solution must wait for further experimentation and study.

Our groups may be compared in yet another way. If final speed indicates the degree of mastery of the maze, we may give the average final speed attained in the retention test as an indication of the ultimate efficiency of the various groups. These data are included in Table 10. The final speed in both learning and relearning, based on the average time per criterial trial, is shown.

TABLE 10

EFFICIENCY OF THE VARIOUS GROUPS

Group	Final speed learning	Final speed relearning	Percentage gain
C	28.75	27.62	3.93
G-20	28.16	25.66	8.88
G-40	36.44	33.75	7.38
G-80	38.33	33.58	12.40

This table shows (*a*) that the guided groups do not reach a final stage of mastery equal to the control, with the exception of G-20, and (*b*) that the guided groups make more progress, in terms of percentage of gain, or improvement, than does the control group. This latter can be accounted for on the grounds that they had further to go than did the control, as judged by their final speed in learning.

SUMMARY AND CONCLUSIONS

The relative efficacy of large amounts of manual guidance, given initially, upon human maze learning has been studied. The significant conclusions may be stated as follows:

1) When the records subsequent to the guidance are considered, by all criteria the efficacy of large amounts of guidance increases with the amount given.

2) The increase in benefit is not in direct proportion but becomes relatively less with the larger amounts. If, from Table 2, we take the average percentage of gain in trials, errors, and time as a rough indication of the influence of guidance, the average gain for

20, 40, and 80 guided trials, or, in terms of time, for 200, 400, and 1600 seconds spent in guidance, is roughly 13, 36, and 46%.

3) Both cul-de-sac and retracing errors are similarly affected, the average decrease for the guided groups in these two criteria are 44 and 49%, respectively. (Table 4.)

4) When guidance is included as a part of the learning records, guidance is detrimental in terms of trials and the deleterious effect increases with the amount given. In terms of time, guidance is first slightly beneficial and then detrimental. The detrimental effect is not, however, as marked in the case of time as it is in trials. (The effect on the error records remains unchanged, of course.)

5) Using final speed as an indication of the degree of mastery, the control group surpasses the guided groups.

6) Guidance operates to reduce the variability of the guided groups in terms of trials and time, and to increase it in terms of errors.

7) On the basis of the saving scores, when subsequent learning is considered, the control group retains the maze habit better than do the guided groups, and there is a tendency for the saving scores to decrease with the amount of guidance given.

8) If saving scores are based on the learning records with guidance included as part of the learning, the saving in trials and time increases slightly with the amount of guidance, and, in terms of trials, those given the larger amounts of guidance retain the habit better than do the controls.

9) In terms of absolute recall, based on the first relearning trial, the guided groups retain more than the control in terms of errors; Groups G-20 and G-40 retain more, Group G-80 less than the control group in terms of time.

10) The ultimate mastery of the maze, based on the final speed in retention, is greater in the case of the control group than in the guided groups, with the exception of G-20, although the guided groups make a larger relative gain in comparison with their final speed in learning.

The chief difference between the present study and earlier ones is to be found in the effect upon the trial and time criteria. There is a rather consistent agreement that, with guidance considered as part of the learning, the larger amounts tend to be detrimental on the trial records. With this our results agree. But when we consider the effect on trials subsequent to guidance, the conclusions vary. In most of the researches the larger amounts of guidance are less beneficial than the smaller. This is contradictory to our results, since we find the opposite to be true. It is also contradictory to the work of Carr, Carr and Osbourn, and Alonzo. In these latter researches the discrepancy can be accounted for on the basis of the fact that their subjects were given a type of guidance which permitted them to take a more active part in learning. This, however, does not account for the difference between our results and others. They find, to repeat, that the effect first increases and then decreases. We find the beneficial effect to increase with the amount. In accounting for this difference, it should be pointed out that the smallest amount of guidance used in this experiment is larger than the largest formerly employed. Possibly the full effect of guidance is obtained only when given in large amounts. At present we are unable to offer any explanation for this difference.

As for the time criterion, as has been mentioned, very little use is made of it by earlier investigations. Three use it. Koch considers it as a part of the total learning time; Alonzo computes the time subsequent to guidance; Wang uses it only as a measure of the final speed in learning. Hence no extended comparison can be made. Our results are in harmony with Alonzo's when time subsequent to guidance is considered. Our results do not agree with Koch's, since only our smallest amount of time spent in guidance was beneficial. Here again, however, the differences in the amount of guidance may account for the differences found. Our results with regard to final speed in learning are roughly in harmony with those of Wang, viz., that guidance is detrimental.

Koch is the only one to investigate the factor of retention, and she utilizes the absolute recall score based on the errors and time for the first run after an interval of 48 hours. Here the chief difference, basing the comparison on her initially guided groups, consists in the relatively larger amounts retained by her groups. This, of course, was to be expected in view of the longer interval in our experiment. The general tendency of her results, however, is the same as ours, viz., the guided groups retain the maze habit, as measured by both errors and time, better than the control, with the exception of the time criterion in our Group G-80 and the group given two guided trials in Koch's experiment. In neither experiment, moreover, is there any relation between the amount of guidance given and the degree of retention.

REFERENCES

1. ALONZO, A. S. The influence of manual guidance upon maze learning. *J. Comp. Psychol.*, 1926, **6**, 143-157.
2. CARR, H. A. The influence of visual guidance in maze learning. *J. Exper. Psychol.*, 1921, **4**, 399-417.
3. CARR, H. A., & OSBOURN, E. B. Influence of vision in acquiring skill. *J. Exper. Psychol.*, 1922, **5**, 301-311.
4. KOCH, H. L. The influence of mechanical guidance upon maze learning. *Psychol. Monog.*, 1923, **32**. Pp. 113.
5. LUDGATE, K. E. The effect of manual guidance upon maze learning. *Psychol. Monog.*, 1924, **33**. Pp. 65.
6. MCGEOCH, J. A., & MELTON, A. W. The comparative retention values of maze habits and of nonsense syllables. *J. Exper. Psychol.*, 1929, **12**, 392-414.
7. WANG, T. L. The influence of tuition in the acquisition of skill. *Psychol. Monog.*, 1924, **33**. Pp. 51.

University of Arkansas
Fayetteville, Arkansas

L'INFLUENCE D'UNE GRANDE QUANTITÉ DE DIRECTION MANUELLE SUR L'APPRENTISSAGE DU LABYRINTHE PAR LES ÊTRES HUMAINS

(Résumé)

On a fait cette étude dans le but de déterminer l'influence d'une grande quantité de direction manuelle, donnée au commencement, sur l'apprentissage du labyrinthe par les êtres humains. On a donné 0, 20, 40, et 80 épreuves, dirigées au commencement, à quatre groupes de sujets. On a donné des tests de rétention après un intervalle de sept jours. Les conclusions signifiantes suivent:

1. Quand on considère les résultats subséquents à la direction, selon toutes les critères l'efficacité d'une grande quantité de direction s'accroît avec la quantité donnée.

2. L'accroissement en valeur n'est pas en proportion directe mais devient relativement moindre avec les plus grandes quantités.

3. Quand la direction est incluse comme partie des résultats de l'apprentissage, la direction est nuisible en termes d'épreuves et l'effet nuisible s'accroît avec la quantité donnée. En termes de temps, la direction est d'abord un peu profitable et ensuite nuisible.

4. Quand la vitesse finale est employée comme indication du degré de maîtrise, le groupe de contrôle dépasse les groupes dirigés.

5. Sur la base des "saving scores", quand on considère l'apprentissage subséquent, le groupe de contrôle retient l'habitude du labyrinthe mieux que les groupes dirigés et les "saving scores" tendent à décroître avec la quantité de direction donnée.

6. En termes du rappel absolu, basé sur la première épreuve de nouvel apprentissage, les groupes dirigés retiennent plus que le groupe de contrôle à l'exception du résultat de temps pour le groupe ayant 80 épreuves dirigées.

7. La maîtrise finale du labyrinthe, basée sur la vitesse finale de rétention, est plus grande pour le groupe de contrôle, à l'exception du groupe ayant 20 épreuves dirigées.

WATERS

DER EINFLUSS HÄUFIGER HANDFÜHRUNG AUF DIE MENSCHLICHE ORIENTIERUNG IM LABYRINT

(Referat)

Das Ziel dieser Untersuchung war, festzustellen, welchen Einfluss häufige am Anfang gegebene Handführung auf die menschliche Orientierung im Labyrint hat. Man gab vier Gruppen von Versuchspersonen entweder keine, oder 20, 40, und 80 anfangs geführte Versuche. Nach sieben Tagen gab man Gedächtnisteste (Retention test). Die bedeutungsvollen Schlussfolgerungen sind folgende:

1) Wenn man die der Führung folgenden Beobachtungen in Betracht zieht, so ergibt sich, dass die Wirksamkeit häufiger Führung nach allen Kriterien mit der Quantität der gegebenen Führung zunimmt.

2) Die Zunahme des erwachsenden Vorteils ist nicht genau proportional, sondern nimmt mit grösserer Zahl der Führungen ein wenig ab.

3) Wenn Führung als Teil den Lernaufzeichnungen eingefügt wird, so ist Führung bezüglich der Versuche nachteilig, und die nachteilige Wirkung wächst mit der gegebenen Anzahl der Führungen. Hinsichtlich der Zeit ist Führung zuerst ein wenig vor- und hernach nachteilig.

4) Nimmt man die Endgeschwindigkeit als Ausdruck des Beherrschungsgrades, dann überholt die Kontrollgruppe die geführte.

5) Auf Grund der Einsparquoten (saving scores) ergibt sich, dass wenn das nachfolgende Erlernen berücksichtigt wird, die Kontrollgruppen die Orientierungfähigkeit besser beibehalten als die geführten Gruppen; und die Einsparquoten tendieren mit der zunehmenden Führung abzunehmen.

6) Hinsichtlich absoluten Gedächtnisses, ergibt sich auf Grund der ersten Wiedererlernversuche, dass die geführten Gruppen besser als die Kontrollgruppen im Gedächtnis behalten, mit Ausnahme der Gruppe, die 80 geführte Versuche erhielt, hinsichtlich der Zeit.

7) Die endgültige Beherrschung des Labyrints ist auf Grund der Endgeschwindigkeit bei den Gedächtnisversuchen für die Kontrollgruppe besser, mit Ausnahme der Gruppe, die 20 geführte Versuche erhielt.

WATERS

VISUAL PERCEPTION VERSUS VISUAL PLUS KINAESTHETIC PERCEPTION IN JUDGING COLORED WEIGHTS*[1]

From the Psychological Laboratories of the University of Minnesota

Cornelia D. Taylor

Historical Introduction

The effect of color upon apparent weight has been investigated by two different procedures which may be briefly introduced by two questions. The first question is suggested by the more or less popular belief that darker colors appear heavier than lighter ones. Can it be experimentally demonstrated that such a belief exists among large numbers of people, or even among a majority of those tested? And, secondly, does color exert an influence upon judgments of lifted weights so that we may speak of a color-weight illusion as we speak of a size-weight illusion?

The first attempt to determine experimentally the effect of color upon apparent weight was made in 1905 by Bullough (1). He presented to his observers a series of pairs of triangles and requested them to state which of each of the two figures they preferred, and why they made their choice. The triangles in any pair differed only in one respect, namely, the lower half of one was of the same color as the upper half of the other, and vice versa. The majority of his subjects preferred the triangle with the lower half of a darker color than the upper half, rather than the one in which the colors were reversed. Bullough considers the basis of this choice to be: "Attribution of heaviness to colors of less, and of comparative lightness to colors of greater luminosity."

Using "weight" to mean sensations aroused by lifting colored cubes, and investigating the problem of the apparent weight of colors by several different techniques, DeCamp (2) obtained highly inconsistent and inconclusive results. By the method of paired

---

*Received for publication by Walter R. Miles of the Editorial Board, October 11, 1929.

[1] The writer wishes to acknowledge her indebtedness to Dr. M. A. Tinker who suggested the problem and gave generously of his time and advice; and to Dr. Charles Bird for critical reading of the manuscript.

comparisons, nine colors were found to have the following rank order from heaviest to lightest: *R, W, O, V, G, P, B, Y, Bk.*[2] The following arrangement was also obtained by comparing each colored block with each of a series of grey blocks graded in weight: *V, W, P, Bk, O, B, R, G, Y.* Finally, by the order of merit method of comparison, red was judged the heaviest, yellow next, then green, then blue.

Another series of experiments performed a year later by DeCamp failed to yield results which agreed very closely with any of the arrangements in the first series; nor were they any more consistent.

None of these arrangements conform to the rule that dark colors are judged heavier than light colors, nor do they consistently point to the operation of a color-weight illusion which would cause the dark colors, while appearing heavier, actually to be judged lighter when lifted.

When, however, DeCamp had his subjects rank six colors in order from heaviest to lightest by stating which looked[3] the heaviest, the next heaviest, and so on, the final arrangement was *Bk, B,* (*R* and *W*),[4] *Gy, Y.* This, except for the misplacement of white, is approximately an order of increasing luminosity.

Koch (4) in 1923 investigated the problem by a modification of the "looks" method which involved adjusting a rod pivoted at its center to an elliptical grey background until colored disks at either end of the rod appeared balanced. The colors in order of apparent heaviness (based on Koch's results from six subjects) were *G, Bk, Gy, R, B, W, Y.* By haphazardly moving the bar, Koch determined the following chance order of colors: *R, Gy, G, Y, W, Bk, B.* He points out that consistency is no greater among his six subjects than among the six chance samples. He concludes from this that there is "no uniform selective color-weight factor for the group," but that with some individuals such a factor may operate.

To discover the factors determining the apparent heaviness of colors, Monroe (5) had his subjects adjust a vertical pointer as if it were a fulcrum until colored disks at either end of a horizontal bar appeared balanced. Using pairs of colors of the same hue and saturation, but differing greatly in brightness, he found that the "apparent weight of a color varies inversely with its brightness." When his pairs of colors were of the same saturation and brightness but different in hue, he found that the colors in order from heaviest to lightest were *B, R, G, Y*. In his third series, the pairs of colors were the same in hue and brightness, but very different in saturation. Results from this series show great individual variation, but usually the more saturated color appeared the lighter.

---

[2] Throughout this report the following set of symbols will be used: *R*=red, *O*=orange, *Y*=yellow, *G*=green, *B*=blue, *V*=violet, *P*=purple, *Bk*=black, *Gy*=grey, *W*=white.

[3] This may be termed the looks method in contrast with the "hefting" method employed in the first part of DeCamp's experiment. Bullough (above) used the "looks" method.

[4] Red and white had the same rank.

An experiment on the effect of color on apparent weight and size, by Warden and Flynn (6), attempted to duplicate retail trade conditions by placing in a glass show case eight differently colored cartons identical in shape and size. Observers wrote opposite the name of a color their judgment of its relative weight in each of eight arrangements.

The experimenters found that, while the influence of a color on apparent weight varied widely from one arrangement to another, most colors had a fairly consistent ranking throughout. The colors in order from heaviest to lightest according to average ranks are *Bk, R,* (*P, Gy*), *B, G, Y, W*. This arrangement approximates closely an order of increasing luminosity.

A summary of the experiments reviewed shows that arrangements of colors in order of apparent weight vary from those agreeing perfectly with an ascending order of luminosity, through some approximating pure chance, to a few which suggest a descending order of luminosity. Dividing the results, however, into two groups, those determined by the "looks" method and those determined by actually lifting colored weights, reduces the chaos to some semblance of order. By this first method only one investigator (Koch) obtained inconsistent results and a final arrangement which approximates chance. The other four experiments yielded arrangements of the colors which in order of apparent weight agree closely with a ranking from darkest to brightest. DeCamp was the sole investigator to employ the "hefting" method, and his results vary with the number of colors used and with the specific method of investigation. To obtain a clearer and more definite concept of the influence of color on apparent weight further experimentation is necessary. We need carefully controlled work which will compare and correlate judgments made under different methods by large numbers of subjects.

The present study was planned to answer the following questions:

(*a*) Does color have an influence upon the apparent weight of objects? (*b*) What is the relation between rankings of colors in order of apparent weight obtained by the "looks" method and similar rankings obtained by the "hefting" method? (*c*) What is the reliability of each of these methods? (*d*) What is the relation between rankings obtained by the method of paired comparisons and rankings obtained by the order of merit method? (*e*) Do the sexes rank the colors differently?

Apparatus

The apparatus employed in this experiment consisted of six 2½-inch cubes made of cardboard and covered with colored paper from the Milton-Bradley standard series. The boxes were partially filled with lead shot embedded in cotton so that each box weighed just 60 grams. The colors used were red, yellow, green, blue, black, and white. The subjects were seated before a table covered with a dull black cloth. On three sides of the table were black cardboard screens rising 2½ feet above the table top. On the table about 8 inches from the front edge stood two black iron supports, and stretched between them, 10 inches above the table, a white thread indicated the height to which the cubes were to be lifted.

The experiment was performed in a dark room with a constant artificial illumination of 35 foot-candles as measured by the Edison Foot-Candle Meter.

Experimental Procedure

The experimental procedure may be described under two headings, according to the method of comparison employed. The first of these was the method of paired comparisons. The subjects used were 50 men and 50 women students at the University of Minnesota, most of whom were members of elementary laboratory classes in psychology. After the subject had been seated before the table, he was handed one of the following sets of directions. The first set of instructions is for the "looks" procedure; the second is for the "hefting" procedure.

1) This experiment is to determine how heavy certain objects appear when compared with others.

Two cubes will be placed before you on the table. You are to state which one looks the heavier.

On each comparison the experimenter will state which of the presented cubes is the standard, and you are to state whether the other cube appears heavier or lighter than the standard. You will be allowed about four seconds to make your judgment.

2) This experiment is to determine how well you can discriminate the difference between certain weights. These weights are all different: no two weigh the same.

The second weight in each comparison is to be judged in terms of the first; that is, the second is called lighter or heavier than the first.

The blocks are grasped like this (demonstration) and lifted about 10 inches. Do not let your hand strike the string stretched across the table. The eyes are to be kept fixed upon the block while it is being lifted. Judgment is to be made immediately after lifting the second weight.

For half of the subjects, namely, the last 25 men and the last 25 women, the directions for the "hefting" procedure were modified so that second paragraph read:

The first weight in each comparison is to be judged in terms of the second; that is, the first is to be called lighter or heavier than the second.

The manner of grasping and lifting the cubes was demonstrated by the experimenter, and, except for slight variation in the speed with which they were raised, was constant for all observers. The cubes were placed directly before the subject about 8 inches from the front edge of the table and a little more than the width of a cube apart.

To control possible time and space errors two forms of presentation were devised and used alternately. Thus, by combining results from all subjects, for the total number of comparisons in which any one color was involved, that color appeared half of the time to the left and half of the time to the right; also, half of the time it was the standard, and half of the time the variable.

Twenty-five of the men, every other one in order of testing, employed the "looks" procedure at their first sitting and one week later the "hefting" procedure at the second sitting. The remaining twenty-five men used the "hefting" method preceding the "looks" method. The fifty women were divided in exactly the same way.

At each sitting, each subject was given, preceding the experiment proper, a practice series of three trials in which 2-inch cubes weighing only 30 grams were used as stimuli.

To determine whether judgments of apparent weight are influenced by the methods used, 30 subjects, all women, ranked the cubes by a modified form of the order of merit method. Fifteen of these women employed the "looks" procedure, and 15 the "hefting" procedure. This time all six cubes were placed on the table before the subject in haphazard order. The directions given to the subjects for the "looks" procedure appear below:

1) This is an experiment to determine how heavy certain objects appear when compared with others.

I shall place six cubes on the table before you. I shall select any two at random and you are to tell me which of the two looks the heavier. I shall place these two near the front of the table, the heavier one to the left, the lighter one to the right.

Then I shall pick a third weight at random from among the remaining four and ask you to tell me whether it looks heavier or lighter than the first one. If it looks *heavier,* I shall place it to the left of this first one; if it looks *lighter,* I shall ask you to compare it with the second weight. If it looks *heavier* than this second weight, I shall place it *between* the first and second weights; if it looks *lighter* than this second weight, I shall place it to the right of the second weight so that it becomes the third weight in the series.

We shall continue in this manner until all the weights have been arranged in serial order.

Then you will check your arrangement by comparing each weight again with the one immediately following, asking me to re-arrange them if necessary, so that in your final arrangement the first weight looks heavier than the second, the second looks heavier than the third, the third looks heavier than the fourth, and so on.

Directions for the "hefting" procedure were worded similarly except that they instructed the observers themselves to select the cubes to be compared and to actually lift them.

Three of the smaller cubes again comprised a practice series which preceded the experiment proper and served to clear up any questions as to the exact procedure desired.

Results and Discussion

In the method of paired comparisons each of the six colors was compared once with every other color. Thus, for any one individual, each color could be given a "score" in terms of the number of times it had been judged heavier than the ones with which it had been compared. The maximum score possible was five and the minimum zero. Combining individual records gave an average score for each color. This score represents the relative apparent weight of that color.

Data obtained by the order of merit method of comparison were analyzed in a somewhat similar manner. By this method the six colors were ranked by each subject in order of apparent weight from the heaviest (Rank 1) to the lightest (Rank 6). An average rank was computed for each color and these averages yielded a final arrangement of the six colors in order of apparent weight.

To determine the reliability of the various methods of investigation, the total group using any one method was split, and the final rankings of the colors obtained from one half (Group A) were correlated with similar rankings based on the other half (Group B). The 50 men employed as subjects in the method of paired comparisons were divided so that, in the two groups of 25 each, practice and methods of procedure were equated. The 50 women were divided in the same way. A third check upon the reliability of this method was secured by correlating the combined rankings of Groups A for men and women with similar rankings of Groups B for men and women. For the order of merit method the final rankings of the colors based on the judgments of the first seven subjects were correlated against those based on the judgments of the last eight. All of the rank difference correlations were converted into product moment correlations by means of the formula $r=1.0233\rho$ which gives the most probable value of r for any given value of ρ. The resulting reliability coefficients are presented in Table 1. The obvious conclusion to be drawn from the coefficients shown in the first row of the table is that, irrespective of the sex of the subjects or the method of comparison employed, more consistent results are always yielded by the "looks" procedure than by the "hefting" procedure. Reliability coefficients based on the "looks" procedure are all high, two of them being above .90, one nearly .90, and the lowest +.82. Application of Brown's prophecy formula $r_x=\dfrac{Nr}{1+(N-1)\,r}$ raises all these coefficients to +.90 or above. Such coefficients indicate that the rank orders of the colors obtained by the "looks" technique have been established with a high degree of reliability. We, therefore, are justified in placing considerable emphasis upon results obtained through the use of this method. In the "looks" procedure men are slightly more reliable than women, while in the "hefting" procedure they are very much more reliable. The lowest of the four reliability coefficients based on the "looks" procedure is that in which the order of merit method of comparison was employed. A logical explanation of this fact is that the order of merit method, even as modified in this experiment, is more complicated than the method of paired comparisons, since presenting all six colors simultaneously makes the isolation and comparison of any two very difficult.

TABLE 1

RELIABILITIES WITH WHICH THE RANK ORDERS OF THE COLORS HAVE BEEN ESTABLISHED BY THE VARIOUS PROCEDURES

(r is computed from Pearson's product moment formula, r_x from Brown's prophecy formula.)

	Method of paired comparisons						Order of merit method of comparison	
Coef-	Men		Women		Men and women			
ficient	Looks	Hefting	Looks	Hefting	Looks	Hefting	Looks	Hefting
r	+.936	+.775	+.899	+.178	+.936	+.436	+.823	+.618
r_x	+.967	+.873	+.947	+.302	+.967	+.607	+.903	+.764

The "hefting" procedure is much less reliable than the "looks" procedure. Here the highest coefficient (r) is .77, which is raised by the Brown prophecy formula to .87. Three of these reliability coefficients are high enough to make group predictions reasonably safe.[5] The reliability coefficient based on the judgments of women for the method of paired comparisons shows a relationship very little better than chance. For the "hefting" procedure, the order of merit method yields a reliability coefficient which is considerably higher than either of those based upon the method of paired comparisons in which women are subjects. Two possible explanations may be considered. In the first place, because of the few colors used (only six) the probable errors of some of these correlations (all except those in the nineties) are so large that they cannot be accepted at face value as "true" reliability coefficients. In the second place, actually lifting the cubes may have more or less effectively isolated from the other four the two cubes to be compared and thus have partially lessened a source of confusion which lowered the reliability of this method under the "looks" procedure. In summary, we note that except for results based on women for the "hefting" procedure in the paired comparisons method, the reliability coefficients, either r or r_x, are high enough to justify group comparisons.

---

[5] Truman L. Kelley, cited by Hunter (3), states that a measuring instrument should have a reliability of .90 or better if the results are to be used for individual comparisons; a reliability of .40 or better if group differences are to be investigated.

Another approach to the reliability of the various methods employed may be gained by computing the coefficients of reliability of the differences between the average scores for the individual colors. Treating the "looks" procedure first, we find that of 15 possible differences between the six averages based on judgments of both men and women only one may be considered statistically insignificant. The smallest of the remaining 14 differences is 2.8 times its sigma, indicating that, upon a repetition of the experiment under similar conditions, the chances of obtaining a difference in the same direction are 99.74 out of 100. When we consider the averaged judgments of either of the sexes we find that for the men, 13 of the 15 differences are greater than 2.8 times their standard deviations and for the women 12 of the 15 are highly significant statistically.

The "hefting" procedure gives smaller differences between the averages of those based on both sexes together or upon women alone, only 4 of the possible 15 differences are of comparatively high significance. Of the results secured from the men alone, however, 8 of the 15 differences are quite reliably established.

We note, therefore, that the two methods of determining reliability agree in pointing out (*a*) for the method of paired comparisons the "looks" procedure is more reliable than the "hefting" procedure, (*b*) using the "looks" procedure men are slightly more reliable than women, and (*c*) with the "hefting" procedure men are considerably more reliable than women.

Table 2 presents the average number of judgments of "heavier" for each color according to the method of paired comparisons. The colors in the first column are listed in order of luminosity from the darkest (black) to the brightest (white). Inspection of the other columns shows that both sexes have a wider range in averages for the "looks" procedure than for the "hefting." The widest range of all is found with the "looks" procedure for men where the color appearing heaviest (black) has an average "score" of 3.98 and the color appearing lightest (white) has an average "score" of 1.94.

For the "looks" procedure, when the scores of both sexes are combined, the only difference of little or no significance is that between the average score for red and the average score for blue. Treating the sexes separately leaves, of the possible 15 differences, 13 for the men and 12 for the women which are quite reliably established.

Ranking the colors in order of apparent heaviness, according to

TABLE 2

THE AVERAGE NUMBER OF JUDGMENTS OF "HEAVIER" FOR EACH COLOR. METHOD OF PAIRED COMPARISONS

One hundred subjects: 50 men and 50 women

Colors	Men				Women				Men and women			
	Looks		Hefting		Looks		Hefting		Looks		Hefting	
	Mean	$\sigma_{dist.}$	Mean	$\sigma_{dist.}$	Mean	$\sigma_{dist.}$	Mean	$\sigma_{dist.}$	Mean	$\sigma_{dist.}$	Mean	$\sigma_{dist.}$
Black	3.98	1.58	3.14	1.12	3.80	1.47	3.28	0.87	3.89	1.53	3.20	1.09
Red	3.34	1.14	3.32	0.84	3.42	1.05	2.86	1.09	3.38	1.10	3.09	1.00
Blue	3.36	1.11	2.38	0.95	3.38	1.16	2.44	1.07	3.37	1.14	2.41	1.01
Green	2.74	1.10	3.34	1.17	2.70	1.02	3.14	0.97	2.72	1.06	3.24	1.08
Yellow	2.64	1.20	2.52	0.93	2.60	1.27	3.18	1.33	2.62	1.23	2.83	1.20
White	1.94	1.45	3.34	0.92	1.98	1.25	3.08	0.87	1.96	1.35	3.21	0.91

these averages, gives a clearer idea of the influence of color upon apparent weight. Inspection of Table 3 reveals a high degree of consistency among rankings of the colors based upon the "looks" procedure by the various groups. We observe also that, in two of these arrangements (Columns 4 and 6), the correlation between luminosity and apparent weight is perfect and negative, the darkest colors appearing the heaviest and the brightest appearing lightest. With this procedure men and women made similar rankings of the colors.

No such consistency exists, however, among the rankings based on the "hefting" method. Red and black occupy different places in each of the three rankings; yellow holds the same position in two but a much different one in the third ranking; while blue is the only color to maintain the same position in all three rankings.

TABLE 3

COLORS RANKED IN ORDER OF APPARENT HEAVINESS

One hundred subjects: 50 men and 50 women

Rank	Men		Women		Men and women	
	Looks	Hefting	Looks	Hefting	Looks	Hefting
1	Bk	G*	Bk	Bk	Bk	G
2	B	W*	R	Y	R	W
3	R	R	B	G	B	Bk
4	G	Bk	G	W	G	R
5	Y	Y	Y	R	Y	Y
6	W	B	W	B	W	B

*The same rank.

Such shifting about in ranks is not surprising when we realize that, for both sexes together, only 4 of the 15 differences between averages are statistically significant. The four significant differences designate blue, a color of relatively low luminosity, as definitely lighter in apparent weight than any of the other five colors except yellow. Similar data from either sex alone show the same tendency to give blue an average score significantly lower than that of any of the other colors save red for the women and yellow for the men.

Table 4 reveals that the correlation between the rankings obtained from men and women for the "hefting" procedure is only .28. The interpretation that this low correlation is due more to the low reliability of the "hefting" procedure than to any real sex differences seems to be substantiated by the high correlation (.94) between the ranking secured from men and that secured from women for the "looks" procedure where color, and color alone, could have been the differentiating factor in judgments of apparent weight. A second factor lowering this correlation between the results of the men and those of the women for the "hefting" procedure may well be that for men there tends to be a negative correlation between the "looks" and the "hefting" procedure, while the corresponding correlation is approximately zero for the women.

The correlation of .94 between the rankings of men and of women when the "looks" procedure was used, since it demonstrates convincingly the absence of sex differences in the effect of color upon apparent weight justifies combining the sexes into one large group to give the "truest" ranking of the colors in order of apparent weight.

Theoretically, if the darker colors appear heavier than the brighter

TABLE 4

INTERCORRELATIONS BETWEEN THE VARIOUS METHODS EMPLOYED AND BETWEEN THE RANKINGS FOR MEN AND THOSE FOR WOMEN

Method	Variables correlated	*r*	*P.E.*
Paired comparison	Looks *vs.* hefting—(men)	—.407	.24
	Looks *vs.* hefting—(women)	.086	.29
	Looks *vs.* hefting—(men and women)	—.200	.28
	Men *vs.* women—(looks)	.943	.03
	Men *vs.* women—(hefting)	.281	.26
Order of merit	Looks *vs.* hefting	.322	.26
	Paired comparison *vs.* order of merit (looks)	.658	.16
	Paired comparison *vs.* order of merit (hefting	.329	.26

ones, they should, when lifted, also feel lighter, since the unconscious tendency to put more effort into lifting objects which appear heavier results in their being raised more easily. Such a color-weight illusion, had it operated consistently for the subjects in this experiment, would have given negative correlations between rankings based on the "looks" procedure and rankings based on the "hefting" procedure. The first correlations in Table 4, then, looks promising, but its probable error is so large that only a series of correlations all negative and all of about this same magnitude could establish it as a "true" correlation (i.e., likely to be obtained upon a repetition of the experiment under similar conditions). The next correlation, however, between the "looks" and the "hefting" procedure with the 50 women as subjects is approximately zero, and the third, based on all subjects, is about midway between the first and the second. These low correlations, with large probable errors, due largely to the few colors used, make it impossible to predict their size upon repetition of the experiment. Also, the unreliability of the "hefting" method, especially with women, is probably an important factor in reducing the size of the coefficients. A most probable cause of this unreliability, and a direct cause of the lack of any significant correlation between results in the "looks" and the "hefting" procedure, is the factor of inadequate control of attitude in the subjects during the "hefting" situation. Instructing the subjects to compare the cubes does not prevent complete disregard of their colors, nor does it eliminate pressure sensations in the finger tips and kinaesthetic sensations in the wrist and forearm as the sole bases of differentiation. Since the cubes were objectively identical in weight, it is not surprising that such behavior should lead to equivocal results bearing no relation to color. There were several persons whose spontaneous remarks at the close of the experiment indicated that they had been consciously trying to disregard the colors in making their judgments. There were other observers who always judged the cube lifted second as heavier than the one lifted first. Since the light and dark colors were equally divided with respect to time sequence, such responses must mean that these individuals were trying to eliminate the colors as factors in their judgments.

Such individuals might or might not, in the "looks" procedure, have adhered to the principle that colors of lower luminosity appear heavier than those of higher luminosity. As an example of the former we may cite results from one subject for whom, by the "looks" procedure, black appeared the heaviest color, and next in order came red, blue, green, yellow, and white. By the "hefting" procedure, however, white, green, and red each had a "score" of two; and blue, yellow, and black had scores of three. There were several individuals with responses very similar to this and many more who approximated it closely. On the other hand, there were some observers for whom the colors in order of apparent heaviness ranged from darkest to lightest (*Bk, R, B, G, Y, W*) in the "looks" procedure and approximated a reverse arrangement in the "hefting" procedure [i.e., *R, Y,* (*W* and *G*), (*Bk* and *B*)]. Such persons, we may assume, had interpreted the instructions for the "hefting" procedure as they were meant to be interpreted if color were affecting their judgments; they were committing the stimulus error and were influenced by a color-weight illusion.

Results from the order of merit method of comparison are presented in Table 5 which gives the average rank of each color for both the "looks" and the "hefting" procedures. The differences between the averages are slight and the standard deviations are large. As in the method of paired comparisons there is a much wider range in the averages based on the "looks" procedure than in those based on the "hefting" procedure.

Re-ranking the colors from heaviest to lightest according to the "looks" procedure yields the following arrangement: *B, Bk, R, W, G, Y*. A similar arrangement for the "hefting" procedure is *Bk, R, W, G,* (*B* and *Y*). In Table 4 we find that the correlation between these two arrangements is .32 with a probable error of .26. This positive correlation is not due to a correspondence between rankings based on the "hefting" procedure and an order of increasing

TABLE 5

AVERAGE RANK OF EACH COLOR DETERMINED BY THE ORDER OF MERIT METHOD OF COMPARISON

Thirty women subjects

Color	Looks		Hefting	
	Mean	$\sigma_{dist.}$	Mean	$\sigma_{dist.}$
Black	3.30	1.83	3.03	1.50
Red	3.90	1.71	3.50	1.41
Blue	2.63	0.34	4.70	0.46
Green	4.63	1.14	4.50	1.55
Yellow	4.97	1.45	4.70	1.72
White	4.57	1.57	3.57	1.69

luminosity. On the contrary, it is explained by the fact that rankings yielded by the "looks" procedure, in the order of merit method of comparison, do not agree with an order of increasing brightness. If black and white are removed from the arrangement based on the "looks" procedure, the remaining colors stand in an order of increasing luminosity. Remarks from several subjects indicated that judgments of the relative apparent weight of black and white were rendered uncertain by two contradictory associations. The first was that darker colors appear heavier than brighter ones; the second, that white clothing makes people look larger than black clothing. Though we cannot be certain that these factors were influencing any but the two or three subjects who voiced them, it is not uncommon for girls to think of colors in relation to dress, and the assumption that such a factor did operate to complicate this method is not altogether unwarranted. The rank difference correlation between arrangements of the colors by the method of paired comparisons and by the order of merit method for the "looks" procedure is +.658± .16 (from Table 4). Evidently, even such a reliable technique as the "looks" procedure yields results which vary considerably according to the method of comparison employed. This low correlation is probably due to the misplacement of black and white in the arrangement determined by the order of merit method. The correlation between the two methods for the "hefting" procedure is much lower, namely, +.329±.26 (from Table 5). This we might expect from the low reliability of the "hefting" procedure and the few subjects employed in the order of merit method of comparison.

In this experiment, rankings of the colors in order of apparent weight based on the "looks" procedure agree closely with rankings obtained by DeCamp, and Warden and Flynn, whose apparatus was very similar to ours, and by Monroe, whose apparatus, though somewhat different, involved the same principle, (i.e., the "looks" method). Only Koch obtained results, by the "looks" procedure, which fail to agree with the rule, first stated and tentatively proved by Bullough, that dark colors or shades appear heavier than light ones. His results do not agree with those of the present study.

The unreliability of the "hefting" procedure may well be the explanation of the inconsistency among the various rankings of the colors found by DeCamp who used this method.

The present experiment, involving a highly reliable technique (the "looks" procedure with the method of paired comparisons) with a large number of subjects, demonstrates conclusively that color, the only respect in which the cubes differed, was a potent factor in determining their apparent weight, and moreover, that the color with the greatest apparent weight was the darkest color (black), the others following in order of increasing brightness. The "hefting" procedure proved ill-adapted to the investigation of the effect of color on apparent weight since it yielded highly inconsistent rankings. It is possible that a more adequate control of the subjects' attitudes which might insure their committing the stimulus error would demonstrate the operation of a color-weight illusion and result in an arrangement of the colors in which black appeared to be the lightest and white the heaviest color.

Summary and Conclusions

1) A review of the experimental literature revealed somewhat inconsistent and inconclusive results and indicated an approach to the problem which would isolate and measure the relative effectiveness of various methods of investigation in determining the nature and extent of the influence of color upon judgments of apparent weight.

2) The apparatus we used consisted of six 2½-inch cubes weighing 60 grams each and covered respectively with black, blue, red, green, yellow, and white paper. A group of 100 university students, equally divided as to sex, judged the blocks by the method of paired comparisons. Each subject used both the "looks" and the "hefting" procedures. Thirty of the 50 women served as subjects using the order of merit method of comparison, 15 judging by the "looks" procedure and 15 by the "hefting" procedure.

3) Results from the method of paired comparisons are stated in terms of the average number of times any color was judged heavier than those with which it was compared. Results obtained by the order of merit method of comparison are in terms of the average rank assigned to each color.

4) To determine the effect of color upon apparent weight, the method of paired comparisons is usually more reliable than the order of merit method. The "looks" procedure, a highly reliable technique, yields consistent results wherever employed. The "hefting" procedure, in the method of paired comparisons, is definitely unreliable for women; and for men it is consistent enough to justify only group predictions. In the order of merit method of comparison the "hefting" procedure is also only reliable enough to justify group comparisons.

5) By the method of paired comparisons and the "looks" procedure the colors ranked according to apparent heaviness (i.e., according to their average scores based on the total group of 100 subjects) stand in the order *Bk, R, B, G, Y, W*. This arrangement correlates perfectly with an order of increasing brightness, that is, the darker colors appeared heavier. Differences between the average scores of any two consecutive colors in this rank order are, with the exception of the difference between red and blue, statistically significant.

6) The "hefting" procedure in the method of paired comparisons yielded fewer reasonably reliable differences. The men tended to judge the dark colors as lighter than the brighter colors. Judgments of the women, however, showed no consistent relation between the luminosity of the colors and their apparent weight.

7) Rankings of colors based on the order of merit method correlate with similar rankings based on the method of paired comparisons only very slightly when subjects used the "hefting" procedure and only well enough to warrant group predictions when the "looks" procedure was employed.

8) Within the method of paired comparisons, the correlation between rankings of the colors based on the "looks" procedure and those based on the "hefting" procedure was —.40 for the men and approximately zero for the women. A perfect negative correlation between these two procedures would have indicated the operation of a color-weight illusion, the colors which appeared heaviest feeling lightest when lifted and vice versa. Such an illusion seems to have operated to a slight degree with men; but for women the unreliability of the "hefting" procedure was great enough to obscure any relation which may have existed between the apparent weight of the colors and the effect of this on judgments of relative weight.

These results suggest that, with more perfect control of attitude, the operation of a color-weight illusion might have been demonstrated.

9) The largest and most obvious sex difference is a marked tendency toward greater inconsistency among women than among men. For the most reliable of the techniques employed, the "looks" procedure in the method of paired comparisons, men and women agree closely in the average score assigned to any color and almost perfectly in the rankings of the colors in order of apparent weight. For both sexes, the "looks" procedure is more reliable than the "hefting,"

though the "hefting" procedure itself is more reliable for men than for women.

10) Finally, we may draw the following conclusions:

a) The specific results are in part a function of the method employed.

b) The actual lifting of colored objects, identical in size, shape, and weight, by a group of subjects composed of men and women, yields differences in apparent weight which are inconsistent and statistically unreliable, and which bear only a slight positive relation to the luminosities of the colors involved.

c) When colored objects, identical in shape and size, are looked at and not touched, their apparent weight is determined by, and stands in a perfect inverse relation to, the increasing luminosities of the colors employed.

REFERENCES

1. BULLOUGH, E. On the apparent heaviness of colors. *Brit. J. Psychol.,* 1907, **2**, 111-152.
2. DECAMP, J. E. The influence of color on apparent weight. *J. Exper. Psychol.,* 1917, **2**, 347-370.
3. HUNTER, W. S. A reply to Professor Carr on "the reliability of the maze experiment." *J. Comp. Psychol.,* 1926, **6**, 393-398.
4. KOCH, B. C. Color weight. (Master's thesis.) Ohio State University Library, 1923.
5. MONROE, M. The apparent weight of color and correlated phenomena. *Amer. J. Psychol.,* 1925, **36**, 192-206.
6. WARDEN, C. J., & FLYNN, E. L. The effect of color on apparent size and weight. *Amer. J. Psychol.,* 1926, **37**, 398-401.

University of Minnesota
Minneapolis, Minnesota

LA PERCEPTION VISUELLE ET LA PERCEPTION VISUELLE PLUS LA PERCEPTION KINESTHÉSIQUE DANS LE JUGEMENT DES POIDS COLORÉS

(Résumé)

Pour déterminer les poids apparents relatifs des couleurs, six cubes de carton (de même grandeur et poids, couverts respectivement de papier rouge, jaune, vert, bleu, noir, et blanc) ont été présentés à chacun de 100 étudiants universitaires selon la méthode de comparaisons en paires. Chaque sujet a employé deux procédés differents. Dans le Procédé I, il a levé les deux cubes d'une paire l'un immédiatement après l'autre disant lequel a semblé le plus lourd. Dans le Procédé II, il n'a fait que dire lequel des cubes de chaque paire a semblé le plus lourd. La combinaison des résultats

de tous les sujets a donné un poids apparent pour chaque couleur en termes du nombre moyen des fois que la couleur a été jugée plus lourde que celles avec lesquelles elle a été comparée.

Le Procédé I, employant la perception visuelle et la perception kinesthésique, s'est montré une technique absolument inconstante donnant des différences de poids apparent lesquelles sont inconstantes statistiquement et n'ont qu'un petit rapport positif avec les luminosités relatives des couleurs respectives.

Le Procédé II, employant la perception visuelle seule, s'est montré une technique três constante donnant des différences de poids apparent lesquelles sont statistiquement significatives. Plaçant les couleurs en ordre de poids, commençant par le plus lourd, sur la base de ce procédé, nous avons: noir, rouge, bleu, vert, jaune, blanc. Cet ordre semble déterminé par les luminosités relatives des couleurs employées et y a un rapport parfaitement inverse.

Taylor

VISUELLE WAHRNEHMUNGEN GEGEN VISUELLE PLUS KINÄSTHETISCHE WAHRNEHMUNGEN BEIM BEURTEILEN VON VERSCHIEDENFARBIGEN GEWICHTEN

(Referat)

Um das relative scheinbare Gewicht der Farben zu bestimmen, wurden sechs Pappenwürfel (identisch in Grösse und Gewicht, mit rotem, gelbem, grünem, blauem, schwarzem und weissem Papier verschieden überzogen) hundert Universitätsstudenten unter der Methode gepaarter Vergleichung gegeben. Jede Versuchsperson unterwarf sich zwei verschiedenen Verfahren. Im Verfahren I hob sie nacheinander zwei Würfel irgend einer Kombination und berichtete, welcher der schwerere sei. Im Verfahren II meldete sie bloss, welcher Würfel jedes Paares als der schwerere erscheine. Die summierten Ergebnisse aller Gegenastände ergaben ein für jede Farbe scheinbares Gewicht und zwar ausgedrückt in der durchschnittlichen Anzahl von Malen, in denen eine Farbe als schwerer befunden wurde als die mit ihr verglichenen.

Das Verfahren I, das visuelle und kinästhetische Wahrnehmungen einschliesst, erwies sich bestimmt als unzuverlässige Technik, indem sich Unterschiede im scheinbaren Gewicht ergaben, die statistisch unzuverlässig sind und nur geringfügige Beziehungen zur relativen Helligkeit der einzelnen Farben haben.

Das Verfahren II, das nur von visuellen Wahrnehmungen Gebrauch macht, hat sich als höchst zuverlässige Technik erwiesen, die Unterschiede im scheinbaren Gewicht ergibt, die statistisch von Bedeutung sind. Die Rangordnung der Farben von der schwersten zur leichtesten ist nach den Ergenissen dieses Verfahrens: schwarz, rot, blau, grün, gelb, weiss. Diese Ordnung scheint bestimmt zu sein durch die relative Helligkeit der einbezogenen Farben und zwar im Sinn einer umgekehrten Beziehung.

Taylor

LUMINOUS APPEARANCES*

From the Psychological Laboratory of Cornell University

George Kreezer

I. Introduction

Before Katz published his *Erscheinungsweisen der Farben*, psychologists were accustomed to think of a visual experience as being completely described if one merely specified the hue, tint, and chroma of the colors involved. Katz' contribution consisted in pointing out that there are a great many more characteristics in which visual experiences may differ from one another, and in attempting a first classification of the modes of visual appearance that could be distinguished on the basis of such characteristics.[1]

Since Katz' work many studies upon the phenomenology of visual experience have appeared. We may think of them as embracing two separable but related problems. One problem is the classification of all of the various modes of visual appearance. It can be accomplished, obviously, only by a detailed investigation of the characteristics of particular modes of appearance, and by the discovery of common characteristics that shall make possible the comparison and the differentiation of the different modes from one another.[2]

This indicates the nature of the second type of problem—the description of a particular mode of appearance that may on some basis be marked off for investigation. Our present study is concerned with luminous appearances; it represents a problem of the second type. We wish to determine what are the characteristics of luminous appearances which differentiate them from non-luminous appearances,

---

*Received for publication by H. P. Weld of the Editorial Board, August 9, 1929.

[1] Hering, it is true, had earlier indicated certain of these characteristics, but Katz was the first to attempt a systematic survey of all the varieties of visual appearance that may occur.

[2] A mode of appearance, as distinguished from its phenomenological characteristics, is the experience as it comes to us as a total, and which we can represent accurately in words only by the specification of the characteristics which can be observed in it. These different phenomenological characteristics can be conceived of as different aspects of the total experience.

and what different types of luminous appearance may be distinguished.

Such an investigation is desirable because no one has yet attempted the careful description of a representative group of possibly different kinds of luminous appearance. As we shall later see, Hering's remarks on luminosity are cursory and were made before Katz attempted his classification of all modes of visual appearance. Katz states that his account of luminosity and glow does not pretend to be exhaustive but is given only in order to complete his itemization of the varieties of visual appearance. Stumpf is interested only in the possible relation of luminosity to his schema of the attributes of visual sensation. And Bühler, finally, is concerned with the bearing of his psychophysical theory on the impression of luminosity and not primarily with a description of luminous appearances.

II. Historical Setting

a) *Classification of Modes of Visual Appearance.* We shall include here a brief summary of the attempts at classification made by Katz, Martin, and Bühler in order to indicate the general setting of phenomenological problems in vision and to define certain terms that will be used later in this paper.

Katz (9, pp. 1-39) distinguishes eight modes of appearance of colors: film colors (*Flächenfarben*), surface colors (*Oberflächenfarben*), transparent plane colors (*durchsichtige Flächenfarben*), bulky colors (*Raumfarben*), mirrored colors (*gespiegelte Farben*), luster (*Glanz*), luminosity (*Leuchten*), and glow (*Glühen*). His account of surface colors, film colors, and bulky colors is all that we need consider here.

The spectral color is given as an example of the film color. It is described as lacking a definite localization at an assignable distance from the observer, as being visually penetrable and loose in texture (but not showing a visible filling of space in three dimensions), as possessing a frontal-parallel orientation relative to the observer, as lying in a practically even plane and lacking ridges, grains, or similar configurations, and finally, as being aesthetically pleasing (9, pp. 6-14).

Colored paper is given as illustrative of the surface color. It is said to possess definiteness of localization in front of the observer, to lack the loose texture of the film color, to be capable of taking any orientation to the line of regard, and of showing unevenness of configuration.

Colored liquids and transparent fogs are examples of the bulky color. They are colors "which appear *to fill* a definite space in three dimensions. . . . They show this space-filling attribute in a distinct way only when they are at the same time in a true sense transparent," that is, when one can perceive and recognize objects through them (9, pp. 17f).

Katz states that the eight classes of visual appearance distinguished by him may be related to each other by way of intermediate modes of appearance. But he gives no information as to the nature of these intermediates other than to mention a couple of examples of the intermediates between surface and film color (9, pp. 9, 29).

He suggests, as a methodological device in the investigation of the modes of appearance of colors, the reduction of any of them to a film color. This reduction "can be attained, in all cases, by the insertion of a single or double screen with holes, at an appropriate distance between the eye and the source of light" (9 pp. 36-39).

Martin (12), besides wishing to check up on the conclusions of Katz concerning film, surface, and bulky colors, also tried to determine the nature of their intermediates. She grouped the intermediates between surface and film into four classes: the soft surface, the veiled surface, the fogged surface, and the immersed surface. The change in the successive stages consists in an increased obscuring of the original hard surface (12, pp. 462f).

With the apparatus which she set up for obtaining intermediates between bulk and film (an episcotister in front of a sheet of colored paper), she obtained appearances like those of the series already described. .But in addition, two other stages were found, the substantial haze and the cloudy haze. In the substantial haze, the last vestige of surface is gone, and the color is seen as a thick substantial mist. The cloudy haze, finally, approximates the films obtained by the reduction of surfaces, the difference being that it is more suggestive of tridimensionality (12, pp. 470-471).

Martin, then, has defined for us, not separate series between film and surface and between film and bulk, but rather, a single series, between surface and film color, in which the intermediates may be considered as bulk-like. That is, these intermediate forms are appearances in which there is an extension of the color in depth, and in some of which there is a definite back limit. But none of them can, in strictness, be called a bulky color, for this has been defined as a color showing definite limits in depth, both in front and back. We shall discuss this point at greater length later on in connection with a type of appearance observed among our luminous appearances.

Bühler (2, p. 143) distinguishes two directions of modification of visual objects: space-dependent modifications (comprising condensation planes, bulky colors, and surfaces) and illumination-dependent modifications (comprising luminous and lustrous appearances). He modifies Katz' terminology by using condensation-plane (2, pp. 53-57) for Katz' *Flächenfarbe* and Martin's film color. He also mentions a special type of surface characteristic, the coherence-plane (2, pp. 52f., 60-64), which results from the perceptual unification or connection of previously isolated points, as in the impression of a glassy surface.

b) *Luminous Appearances.* Hering (7) briefly discusses luminosity and glow in his article on "The Space-sense and the Movements of the Eyes." "We see something as luminous or lustrous when it is sensed as brighter than pure white things in the same illumination, and when, under the circumstances, the greater brightness cannot be taken as incident light. Flames and glowing bodies are luminous. The luminous sensation has also something bulky about it, compared with the definitely plane-like colors of objects. A glowing iron is seen as luminous through and through; likewise a flame. If we bring a red glowing iron near the sun, it can, under certain circumstances, appear as a body with a red surface color; this is impossible when it is brighter than the brightest surface color of things about. On the other hand, a glowing iron can be seen as self-luminous even though it is not brighter than a pure white surface color in the same illumination; but then the sensation is bulky,[3] and the red of the iron is seen not only on its surface as a plane color, but it extends into the depth of the iron" (7, p. 575).

What then, according to Hering, is the essential requirement for the impression of luminosity? It is, apparently, that the color should not be a surface color, but, instead, should have something bulky about it. Hering's statement of a sufficient condition for this impression, "something is seen as luminous if it is brighter than pure white things in the same illumination," is somewhat obscure. In the light of the rest of the paragraph, all it can mean is that, for a given condition of illumination, there is an upper limit to the brightness that a surface color may possess, and that colors which are brighter than this are seen as bulky and luminous. A color may, however, lie below this upper brightness-limit of surfaces, and still be seen as luminous. But then it appears, not as a surface color, but as a color with "something bulky about it."

---

[3] The term bulky as used by Hering is not synonymous with its use by Katz and Bühler. They have explicitly defined it as descriptive of a color that occupies a definite extent in depth, while Hering evidently means by it merely a color that gives the impression of greater depth than that of the absolutely flat, plane-like color of an object.

It should be noted that, although Hering makes separate mention of glowing bodies and luminous flames, he does not tell us what the difference between them is.

Katz (9, pp. 27f.) modifies Hering's account of luminous appearances by saying that film colors as well as bulky colors may appear luminous, and that to do so they must surpass their surroundings in brightness. Just how great must be this excess over the brightness of the surroundings he says it is not possible to state definitely, "since the impression of luminosity is not always aroused under the same conditions of light-intensity," but he refers to Hering's law as holding in most cases. He has, however, misinterpreted Hering in the statement of this law. Whereas Hering said that *if* a color is seen as brighter than pure white things it is seen as luminous, Katz has interpreted this to mean that a color *must* be seen as brighter than a white to be regarded as luminous. [Hering's statement is: "Wird etwas heller empfunden als die reinweissen Dinge bei derselben Beleuchtung, und kann den gegebenen Umständen nach die grössere Helle nicht als ein aufgeworfenes Licht aufgefasst werden, so sehen wir es leuchtend oder glänzend" (7, p. 575). Katz' statement is: "Meist gilt das von Hering gegebene Kriterium, dass etwas heller als ein Weiss unter denselben Beleuchtungsverhältnissen sein muss, um als leuchtend angesprochen zu werden" (9, p. 28).]

Katz agrees with Hering in his description of glow but adds that it "is distinguished from the impression of simple luminosity by the fact that we perceive a well-outlined form of the glowing body, while in luminous flames no definite limits can be made out."

Stumpf (16, p. 76) contents himself with saying that luminosity stands for the upper region of the scale of visual intensity. He is concerned with relating it to his schema of the attributes of visual sensation, rather than with a description of luminous appearances.

Bühler's doctrine of luminosity is most clearly stated in his reply to Katz' review of his *Erscheinungsweisen der Farben.* "Katz and I agree, so far as I can see, in the evaluation of Hering's short but pregnant remarks on luminosity. They contain two provisions: first, luminosity is 'surplus light'; secondly, all planes are luminous that would appear brighter than the brightest white under the given conditions of illumination. We go on from Hering: What is meant by 'surplus light' and 'brighter than the brightest white'? It means, first, *that there must be a luminosity limen* (*Leuchtschwelle*), and indeed, a variable limen, depending on the prevailing intensity of the illumination. If we, in some way, determine the magnitude of the illumination factor in the impression, we thereby determine also a maximum value for the radiation emitted by reflection from the illuminated plane. To put it briefly, the plane cannot reflect more than it receives. But if its radiation exceeds this value, than the impression of luminosity arises" (3).

It is apparent that Bühler, like Katz, has misinterpreted Hering. Hering's reference to "something that is brighter than the brightest white thing" is made to imply a brightness limen for luminosity, that is, that a color *must* be brighter than this value to be seen as luminous. This misinterpretation is even more evident in Bühler's book. There, in discussing the dependence of the impression of luminosity on conditions of illumination, he says, "Hering has grasped the essentials of this complicated conditioning of the impression correctly; his statement is, *that something must be brighter than a white under the same conditions of illumination in order to be regarded as luminous*" (2, p. 148).

On the phenomenological side, Bühler describes luminosity and glow as both involving the impression of "surplus light." For him, as for Hering and Katz, glow is a special type of luminous appearance (2, p. 147). But he does not state unambiguously how glow is to be distinguished from other forms of luminous appearance. Luminosity, he says, is "surplus light" appearing as a condensation plane; glow, however, may appear either as a condensation plane or as a bulky color (2, pp. 172, 146). It is not clear, then, how a condensation plane which is seen as glow is different from one that is simply luminous. Bühler does not refer explicitly to Katz' statement that "the impression of glow is distinguished from that of simple luminosity by the fact that we can distinguish a well-outlined form of the glowing body . . . ;" he does say, however, "the outer contours of a glowing body are not sharp; they become sharper only when the dark, plane parts, *e.g.,* bits of ash, increase" (2, p. 147), a statement which contradicts that of Katz.

III. Procedure and Apparatus

As already stated, our plan was to bring together stimulus-objects that would provide a representative collection of luminous appearances. These objects, we felt, should include not only such customary ones as heated bodies and flames, but also substances which are rendered physically self-luminous in other ways, as is the case with phosphorescent substances and in vacuum-tube discharge. Our list of stimulus-objects, therefore, includes representatives of each of these four types: high temperature bodies, flames, phosphorescent substances, and vacuum-tube discharge. Some of them can be varied so as to provide a wide range of brightness.

Except where otherwise indicated, the objects were all exposed behind a screen with a hole, through which the observer looked at the object. The entire experiment was conducted in a dark room. The only illumination provided during an observation came from the stimulus-object; the observer's impression, therefore, was that of a bright visual object upon a dark background, the screen not being visible. It served, however, to cut off any side reflections that might have been produced by the stimulus and so have illuminated other objects in the field of view. During the intervals between separate observations, the room was illuminated by a 50-watt lamp. The exposure screen was constructed so as to permit the size and shape of the hole to vary for different stimuli.

Description of Apparatus and Stimulus-Objects. The exposure screen was constructed by attaching a sheet of heavy gray cardboard to a wooden frame 71 x 56.5 cm. A hole, 25 cm. in diameter, was cut in the center of the cardboard. Cards containing smaller holes could be placed in slots behind the 25-cm. hole. A sheet of cardboard that could be raised and lowered, in front of this hole, served to control the time of exposure. The exposure frame was fastened to a table, 75 cm. high.

Stimulus-Objects. In the following account of our stimulus-objects, the names which are printed in italics are the ones by which we shall refer to these objects throughout the remainder of the paper.

1) *Heated copper sheet.* A 25 cm. square, cut from a sheet of 26-gauge planished copper, was fixed in a vertical position, with the copper side toward the observer, 30 cm. behind the exposure-screen. By means of a Fisher gas burner, fixed horizontally with its mouth about 1 cm. behind the center of the sheet, the sheet was heated until a bright visible area about 12 cm. in diameter appeared. Before being presented for observation it was heated for a minute and a half, and then kept heated during the length of the exposure. The brightness of the center of the visible patch, as measured with a Holophane Lightmeter, was 6 millilamberts.

2) *Heated iron rod.* An iron rod with a cylindrical shaft 6 cm. in diameter and a head 1.3 cm. square and 7 cm. thick was fixed in a horizontal position. The head and an extent of about 8 cm. along the rod was heated with the Fisher burner for 2 minutes and then exposed for observation at a distance of 15 cm. behind the screen. The observation lasted until the rod was no longer visible. The brightness of the head of the rod at the beginning of the exposure was about 10.5 millilamberts.

3) *Tungsten filament.* A 200-watt tungsten lamp was placed directly in front of a vertical frame covered with black velvet which served as a background against which the U-shaped filament of the lamp was seen. The usual exposure frame was not used. The lamp was adjusted on a standard so that its long axis made an angle of about 30° with the vertical and its filament was a distance of about 60 cm. from the floor. A Nichols frame and a tube rheostat were used for controlling the current through the lamp. With a current of .52 amperes, the filament could barely be seen as a dull red. With a current of 1.125 amperes, the filament was seen as a dazzling white. Seven intermediate degrees of the current were also used. Table 1 states the strength of the current and the potential difference between the lamp terminals for each of the separate settings.

Table 1

Current in amperes	.52	.57	.63	.70	.76	.84	.92	1.00	1.125
Potential differences in volts	8.75	11.0	15.5	19.5	24.5	29.7	40.0	46.0	53.7

4) *Welsbach mantle.* An illuminated Welsbach mantle was placed 30 cm. behind the exposure screen.

5) *Cigar and cigarette tips.* A cigarette or cigar, placed in a holder, and with its lighted end pointed slightly upward, was set 5 cm. behind the screen. A stream of air was drawn through the cigarette or cigar, when necessary, by means of a rubber tube connected with an aspirator, which was in turn connected with the compressed air system. A stream of compressed air, emitted from a second rubber tube, was used to keep the tip free of ash.

6) *Flames.* A Bunsen burner, with the air supply turned on just enough for all traces of the yellow color of the flame to disappear, furnished one of our flame stimuli. This we shall refer to as the *blue flame.* An ordinary wax candle was used to provide another flame, which we shall refer to as the *yellow flame.* The flames were placed 30 cm. behind the screen. The descriptions that will be given of the flames refer to the main body of the flame and not to the small dark cone at the base of the flame.

7) *Aurora tube.* This is a partially exhausted cylindrical glass tube, 25 cm. in length and 2.5 cm. in diameter.[4] It lights up upon the passage of an electric discharge. The reddish luminous appearance of the anode end, the so-called positive column, extends about half the length of the tube, while the bluish luminosity at the cathode end is confined within a relatively small extent. Each end of the tube was used as a separate stimulus-object. We shall, therefore, refer to them as the *Aurora tube, anode end* and the *Aurora tube, cathode end.* The tube, held in a horizontal position by two loops of thin wire hung from a wooden rod, was placed directly behind a rectangular hole in the screen, 4 cm. by 8 cm. The long axis of this hole was vertical. The portion of the tube seen through the hole as the *anode-end* stimulus-object was that extending from 6-10 cm. from the end of the tube; that exposed for the *cathode end* extended from 1-5 cm. from the end.

The power for sending an electric discharge across the tube was obtained from an induction coil equipped with a built-in Leyden jar, and connected with a 6-volt storage battery. The voltage built up by the coil was such as to enable it to produce a 40-mm. spark in air between the ends of pointed terminals.

---

[4] This tube (cat. no. F3001) and the Geissler tube which we used were obtained from the Central Scientific Co., Chicago, Ill.

8) *Geissler tube.* A Geissler tube, 20 cm. long, with a central elliptical bulb, 6 cm. long, and spherical end bulbs, 3 cm. in diameter, was connected with the induction coil. A discharge through the tube filled the cathode bulb with a blue light and produced in the central bulb, which was made of a fluorescent glass, a bright green color. Through the center of this central green bulb, a pink horizontal band could be perceived. We chose the cathode bulb and the pink horizontal band as our stimulus-objects. We shall refer to them as the *Geissler tube, cathode bulb* and the *Geissler tube, central band,* respectively. These portions of the Geissler tube were presented in the same manner as the Aurora tube.

9) *Luminous paint.* As representative of phosphorescent substances, we selected Devoe's luminous paint and a green phosphorescent powder. With the paint, a uniform circular patch, 16 cm. in diameter, was painted on a piece of stiff drawing paper. To decrease the sharpness of the boundaries of the patch, we laid over it two layers of cheese cloth, one with a circular hole 11 cm. in diameter, and the other with a hole 14 cm. in diameter. The resultant stimulus-object consisted of a patch of paint, 11 cm. in diameter, which was directly visible; an adjoining circular band, about 1.5 cm. wide, covered by one thickness of cheese cloth; and then another band, 1 cm. wide, covered by two thicknesses of cheese cloth. This patch was illuminated by a 200-watt lamp, at a distance of 10 cm., for 5 minutes, and then exposed to the observer from a distance of 30 cm. behind the screen. It was found necessary for each of the phosphorescent substances to be presented immediately after the cessation of the illumination, since its appearance changed very quickly after this moment.

10) *Phosphorescent powder.* For this stimulus we used .6 g. of a green luminous paint powder.[5] The powder was poured out in a small pile 35 cm. behind the screen. Above the pile of powder and opposite the exposure hole was fixed a small mirror, tilted at an angle that permitted an observer seated in front of the screen to see the image of the powder. Before each observation the powder was directly exposed to the light from the 200-watt lamp for 5 minutes. The powder was then presented for observation for 5 seconds.

---

[5] The trade name of the powder is *green luminous paint.* It was obtained from the Radium Limited, U. S. A., New York, N. Y.

With all these stimulus-objects, the observer was seated 1.5 m. in front of the exposure frame. For all except the vacuum tubes, a circular hole, 16 cm. in diameter, was used in the exposure frame.

Except where otherwise noted in our description of the individual stimulus-objects, the observer was permitted to observe the object as long as he wished in any given observation. He was also allowed to report during the observation, except in the case of the phosphorescent objects, the heated rod, and the vacuum tubes. In these cases, he reported immediately after the completion of the exposure.

In addition to the regular series of experiments with the objects just listed, we also performed, from time to time, special experiments for checking up on points that were doubtful. Whenever we mention such experiments, we shall state at the same time the nature of the apparatus and the experimental conditions.

Instructions. Except for the heated iron rod and the luminous paint, the instructions were the same in the case of all the stimulus-objects. They read:

"You will be given a visual experience. Describe the experience as completely as you can, and state the name you would use to characterize the experience as a whole."

For the heated iron rod and the luminous paint, the instructions were: "You will be given a visual experience. Observe the experience throughout its course and then describe this course as completely as you can. Also, state the names you would use to characterize the total experience at any of the stages in this course."

Observers. The observers were Dr. L. B. Hoisington (H), Assistant Professor of Psychology; Dr. H. G. Bishop (B), Assistant Professor of Psychology at Wittenberg College; Dr. F. L. Bixby (Bx) and Dr. S. Feldman (F), Instructors in Psychology; Miss E. F. Wells (W) and Mrs. F. W. Hazzard (Ha), graduate students in psychology; and Mr. G. Kreezer (K), the writer of the present paper.

H, F, W, and K observed throughout the entire course of the experiment. Bx and Ha observed in connection with the *heated copper sheet,* the *heated iron bar,* and the *tungsten filament;* and B in connection with the *copper sheet* and the *iron bar.*

IV. The Principal Characteristics of the Visual Objects

In this section we define the characteristics of our visual objects on which we shall base our definition of the luminous appearance and its varieties. These characteristics are the depth-filling character of the visual object, and the light-dark qualitative integration.

a) *The Depth-filling Character of the Visual Objects.* It is under this heading that the differentiation of colors into films, surfaces, and bulks may be placed. As we shall see, scarcely any of the color-appearances furnished by our stimuli show definite limits in depth that would warrant their being classed either as surface colors or bulky colors. And yet, the variations exhibited by them are so pronounced as to make it questionable whether they can all be lumped together as film colors, as this term has so far been defined. Among these appearances, we have found two main types. First, there are colors which appear volumic, which seem to occupy an extent in depth, without there being at the same time any definite front or back limits to this extent. Secondly, there are colors which appear flat, which do not seem to have any extent in depth, and yet are not seen as definite, hard surfaces. Within both of these classes, variations occur which suggest series running between the extremes of dense and loose. That is, color appearances ranging between the extremes of dense and loose may be either volumic, suggestive of depth, or flat, non-depthlike. In the first class, that of the volumic appearance, our stimuli provided experiences representative of the entire range of variation; in the second class, however, only cases illustrative of isolated points in the series occur. One further point should be noted. This distinction of two types of appearance, a volumic and a flat, does not imply a corresponding classification of the stimuli. There is not always a one-to-one correlation between stimulus and experience. Certain stimuli resulted in volumic appearances for some observers, and flat experiences for others. In most cases, however, a given stimulus-object was apprehended in the same way by all observers.

b) *Qualitative Integration.* A phenomenon observed in connection with many of the appearances is that which we term a light-dark qualitative integration. The visual object as a whole stands out as bright against a darker background. But its quality is not absolutely homogeneous or continuous. Throughout the extent of the visual object a relatively lighter and a relatively darker quality are distributed, but not as discrete bits spatially delimited and sharply distinct from each other. They are merely given together in the experience.

In terms of this characteristic, two different directions of variation may be distinguished. When the light-dark integrations are volumic in character, there may be differences in the relative localization in depth of the light and dark aspects. These differences are, again, of two kinds. First, there are integrations in which the brighter quality is seen as lying or extending further back in depth than the darker quality; the experience seems to get brighter as one penetrates more deeply into it, and a pure homogeneous brightness is taken as lying behind. An iron rod heated to a bright red furnishes a good example of such an appearance. Secondly, there are integrations in which the difference in locality is not observed; both qualities interfuse with each other throughout the entire extent of the experience as, for example, in the luminous discharge of an Aurora vacuum tube.

Besides these differences in the relative localization of the qualities, the observers occasionally specify differences in the dominance of the lighter or darker aspect. In some cases the experience is spoken of as predominantly bright, in others as consisting of the qualities in approximately equal balance. So we may conceive of a series which extends from a color that is uniformly dark to one that is uniformly bright, with intermediates between that vary from integrations with a preponderance of dark at the lower end to those that are overwhelmingly bright at the upper end of the series.

We may now list the stimulus-objects in which the characteristics described above may be observed and, where illustration is considered desirable, give a few samples of our observers' reports.

1) *Depth-filling Character of the Visual Object*

a) *Volumic appearance: dense type*

Stimulus-objects: Welsbach mantle, cigar and cigarette tips, tungsten filament (from about 0.76 A. to 1.1 A.), heated iron rod (within about first 5 sec.), yellow flame, and heated copper sheet.

H: (Welsbach mantle) It is definitely massive, or bulk-like, but there is no transparent surface, so it isn't like a bulky color. *F:* (Iron rod, at beginning) The head was a mass of brightness. It has depth, but is impossible to tell how far the depth extends. (Yellow flame) It has a suggestion of depth or bulkiness, but you can't see through it. It's very bright, compact, and dense. The compact brightness seems to lie behind the yellowish color of the flame. *W:* (Welsbach mantle) It is dense and has a suggestion of thickness or bulk. It isn't flat, but has no definite spread in the third dimension. I can't say where it begins and where it ends. (Tungsten filament, 0.7 A.) There is a narrow strip of hard glow. It is hard, like a surface, but there is no surface there. It has a suggestion of depth, but I can't see into it as I can into bulky colors. (Yellow flame) It isn't a definite bulk, like the bulkiness of colored liquid or ice, yet there is a vague bulkiness about it. It is as if a filmy color were gathered together into a mass. *Bx:* (Heated sheet) The central spot of glow is quite bulky, but not a pure bulk like a bottle of colored liquid. There is no distinct front surface that stops your vision off short. *B:* (Heated copper sheet) It is a motile ball of redness. There is a fluffy part that seems to be further behind. It is that part that does the glowing.

b) *Volumic appearances: loose type*

Stimulus-objects: Aurora tube, anode and cathode ends (*O*'s: H, F, K); Geissler tube, cathode end and central band (*O*'s: H, F, K); blue flame (*O*'s: H, F, K, W) (W described experiences aroused by Aurora and Geissler tubes as *flat* and film-like.)

H: (Aurora tube, cathode end) There is a purplish light interpenetrated with dark. It is definitely depthlike, yet soft and filmy. *F:* (Geissler tube, central band) The pink inside was glow. It was like a bright mist. The texture was loose, but it seemed to harden up inside, also get brighter. W: (blue flame) The whole thing is glow-like. The texture is soft and filmy, but not a flat film. It is like a mass of filmy stuff.

c) *Volumic appearances: intermediate density*

Not all of our stimulus-objects provided appearances that were distinctly dense, or distinctly loose. Some were intermediate in nature.

Stimulus-objects: Appearances covering a considerable range of intermediates were provided by the heated iron rod as it faded out, by the tungsten filament with different magnitudes of current through the lamp, and, except for one observer, by Devoe's luminous paint throughout the course of its fading. The appearance of green phosphorescent powder may be placed at about the middle of the series.

The *heated iron rod* is initially of dense volumic appearance. As it gets darker, it gets looser in texture and hazier in outline, but retains its volumic character almost to the end. Shortly before it disappears, however, it loses its volumic character, its suggestion of depth, and then it is no longer seen as "glowing." The *tungsten filament,* with relatively large currents through the lamp (from 1.1 A. to about 0.76 A.), was seen as dense and massive. But with decreasing currents, it became increasingly loose and soft. It never, however, even with the lowest current (0.52 A.) reached the looseness and filminess exhibited by the Geissler and Aurora tubes. The *luminous paint,* at the first instant of exposure, was seen by H, F, and K as fairly massive and dense. It gradually decreased in brightness and density for about 5 seconds, when it was quite loose and filmy. Then it remained practically constant for a long time. Its volumic character was now considerably less marked, and it was seen as only slightly luminous. W, on the other hand, described her experience with this stimulus-object as a *flat* filmy spread, lacking all bulk or body. She also reported that it was not seen as glowing or luminous.

d) Non-volumic or flat appearances

Stimulus-objects: The final stages of the heated iron rod (all observers), the cathode bulb and the central band of the Geissler tube (W), and the anode and cathode ends of the Aurora tube (W) furnished the stimuli for the loose type of appearance. A range extending from a fairly dense, flat experience to a looser experience was observed by W in connection with the luminous paint. An example of the dense, flat type of appearance was observed by Bx at the initial moment of exposure of the heated iron rod. Some of the stages of the tungsten filament were also occasionally seen by Bx as flat.

H: (Final stage of heated iron rod) There came a time when the total experience showed nothing of a greater brightness behind. The experience was not luminous in any sense, but it's hard to say when, precisely, it ceased to be glow. But when I found that it was not glow, it had lost the back-lying, high-density brightness. It was now all of a piece, a non-surface looseness. *F:* (Final stage of heated iron rod) After a time, the massive brightness behind disappeared altogether. The glow is that massive brightness behind. When that disappeared, there was just this orange plane that was like a film, without discernable thickness. *W:* (Final stage of heated iron rod) As the experience became less bright, it lost its bulky character and became just a flat film. At this time, it was no longer glow, not luminous. (Geissler tube, cathode bulb) There is a spread of filmy blue, a soft, loose-textured, flat film. There is dark gray stuff mixed up with it. This film is not glow-like. *Bx:* (Tungsten filament, 0.57 A.) The left segment doesn't glow at all. It's simply a dead orange color. Taken by itself, it is flat.

2) *Light-dark Qualitative Integration*

a) Variations in relative localization of the qualities: bright quality further back than the dark

Stimulus-objects: Heated copper sheet; cigar and cigarette tips; iron rod as it fades out, before losing its volumic character; tungsten filament (current: 0.57 A. to 0.84 A.); luminous paint or phosphorescent powder at first instant of exposure.

H: (Copper sheet) The whole thing gets brighter toward the center and as you penetrate back into it, till you reach that central portion where it is immensely bright. (Luminous paint, first instant) There is depth to it, and no surface at all. It is lighter further back. *F:* (Cigarette tip) The impression of the total object was that of a massive, depth-like brightness behind, screened by an orangish hue. (Iron rod) Behind the front plane there was an indefinite extent of brightness which thinned as the front became darker. The plane itself was soft, filmy and of orange color. (Tungsten filament, 0.84 A.) An internal mass of brightness seems to lie behind an orange film, but to be much thicker than the film. (Phosphorescent powder) This was glow, particularly the first moment. The bright portions seemed to extend behind the front plane of the visual experience. *W:* (Copper sheet) A grayish film seems to overlay the brighter, denser background. (Iron rod) A gray more or less distinct from the bright reddish orange seems to overlay it.

b) Variations in relative localization of the qualities: distribution of both qualities throughout the extent

Stimulus-objects: Blue flame; Geissler tube, cathode bulb; Aurora bulb, anode and cathode ends; luminous paint, after about first 5 sec. of exposure.

H: (Geissler tube, cathode bulb) There is no suggestion of a brighter back portion. It is more or less uniform all through its suggested depth. The darkness is mixed all through the constitution of the experience. (Aurora tube, anode end) As you go in toward the center, it doesn't ever become pure light; there is always an intermixture of dark. (Luminous paint, end stage) In the beginning the brightness appears as a fairly compact mass lying behind, whereas later it is scattered all through the filmy stuff and isn't a mass at all any more. *F*: (Luminous paint, end stages) It was differentiated into irregularly and indefinitely bounded patches of light and dark. *W*: (Blue flame) A grayishness is mixed up with the experience. The gray and the color are not separate, not spatially distinct. They are beaten up together. (Aurora tube, anode end) There is a film of gray mixed up with the color, but I wasn't able to separate the two. It was as if the darker gray were everywhere and the color were everywhere.

c) Variations in relative dominance of the qualities

Reports upon this direction of variation of the light-dark integration were less frequent than upon the direction of variation just discussed. The stimulus-objects for which the homogeneously bright experience was reported were the tungsten filament (currents of 1 to 1.125 amperes) and the Welsbach mantle. Those for which the predominantly bright integration was reported were the heated copper sheet, the heated iron rod (within about first 15 sec. of exposure), the phosphorescent powder, and the tungsten filament (currents of about 0.76 to 0.92 amperes). Those for which an approximately equal balance of the two qualities was reported were the peripheral parts of the heated copper sheet and the luminous paint (after about 5 sec. of exposure).

Discussion of Results

The Volumic Appearances

Classification. The volumic appearances do not fit very readily into the classifications heretofore proposed. They can qualify neither as surface, film, or bulky colors. Being volumic, they are obviously not surfaces. But neither are they bulky colors. Katz and Bühler have given this term to colors which seem to fill an extent in depth. But they so define their bulky color that one having definite front and back limits is implied. Katz, for example, says, "by bulky colors we mean those colors which appear *to fill* a definite space *in three dimensions*." The reference to a definite space in three dimensions implies that there must be definite boundaries to the color, as in a thin cake of clear ice in which we can definitely see both the front plane, where the color begins, and the back plane, where the color ends. Katz makes this requirement even more specific when he says that a bulky color must "in a true sense be transparent," that is, objects must be visible through it. If this does not occur, as in a very dense fog, "the bulkiness of the whitish color disappears and the white now possesses the character of an indefinitely localized, screening film color" (9, pp. 17f). Bühler (2, pp. 12-15) defines the bulky color in substantially the same way as Katz. When it is impossible to see through a color, he says we must speak of it as a condensation plane, or, in Katz' terminology, a film color. Lacking such definite limits in depth, volumic appearance cannot be considered bulky colors. They qualify just as poorly as film colors. Both Katz and Bühler describe the film color as plane-like (9, p. 8; 2, pp. 53-57), a description not appropriate for the volumic appearance. The dense type possesses the characteristics of film even less than the loose type, for it lacks, besides a plane-like character, the loose, insubstantial texture (9, p. 7; 2, p. 57) and the indefinite localization (2, p. 56) of the film.

An examination of Martin's intermediates between surface and film reveals one stage that approximates the volumic appearance. The "cloudy haze" is described as differing from film in being more positively bulky; "it hinted more strongly at tridimensionality." It is significant that several of her observers referred to this stage as luminous (12, p. 473). Our series of volumic appearances ranging from dense to loose might be conceived of as intersecting at right angles with Martin's series from surface to film color, with the point on her series designated as "cloudy haze" coinciding with the point on our series showing about the same degree of density. Her cloudy haze, then, would be represented as a particular case of our general class of volumic appearances.

The characteristics of reduced colors. In order to confirm our observers' reports of the texture of the loose volumic appearance mentioned above, we set up a film color (reduced from a surface by means of the double-screen method) for purposes of comparison. As a result, we not only made the necessary confirmation but discovered that "reduced colors" of high brightness do not possess the characteristics of the film color. The looseness of texture, penetrability, and indefiniteness of localization do not persist as the color increases in brightness. The reduced color becomes less loose and soft, less penetrable, and more definitely localized; it almost approximates a surface in its smoothness and hardness. This may mean either that colors of high brightness cannot be reduced to film by the traditional double-screen method, or that such bright colors never possess the characteristics which have been defined as those of the film color. We believe the second alternative to be more probable. The efficacy of the double-screen method in the reduction of colors to film has been attributed to its elimination of the usual cues to localization, to the hindrances it imposes upon definite fixation (9, pp. 37f; 12, pp. 464-469). We should expect it to be as effective in this respect with one quality as with another. It would seem, therefore, that the differences discovered in the density and penetrability of the reduced colors are inherent in the qualities themselves. At any rate, this chance observation indicates the need of an investigation of possible phenomenological differences among reduced colors of different quality. Differences may be found among colors of different hue and chroma as well as among those of different brightness.

Our apparatus for the above observations consisted of Katz' double-screen for reducing a surface color to a film color, Martin's arrangement for obtaining intermediates between surface and film (a series of holes of different sizes in the screens) together with a separate 200-watt lamp for varying the brightness of the color. Our surface color was a sheet of cardboard covered with Milton Bradley violet red paper, and placed 45 cm. behind the back screen; the separate illumination lamp was placed 25 cm. in front of the colored paper. A range of 0.9 to 1.8 amperes in the lamp provided the phenomenological variations described above. A 4-cm. hole was used in the screen; a smaller hole was not found to be any more effective in causing the bright reduced colors to assume the characteristics of the film color.

The influence of attitude. A point which we have already mentioned deserves further comment—the lack of a constant relation between the stimulus and the resulting mode of appearance, for it illustrates very prettily a principle emphasized by Katz as being of major importance in the occurrence of the different modes of visual appearance—the influence of attitude. "As compared with the great influence which the inner attitude (*Einstellung*) exercises on the occurrence of this or that mode of appearance of colors, the physical origin of the radiation arousing the impression can be designated as almost irrelevant" (9, p. 9). The mode of appearance that a given color assumes for one will depend, then, not only on the nature of the stimulus, but also on the particular attitude or set of the observer, the state of readiness of his nervous system to be actuated in one way or another by a stimulus. And such a state of readiness should, no doubt, be conceived of much more broadly than as corresponding only to a conscious intention to apprehend an experience in a certain way. We are undoubtedly thrown into all kinds of special attitudes by conditions of which we may be completely unaware. And so, in the present instance, certain stimuli led to volumic modes of appearance for some observers and to non-volumic appearances for others, despite the lack of any corresponding specific instructions.

It is possible that just such a factor as attitude may be the basis of the disagreement among different writers in their characterization of the film color. Katz, for example, refers to this mode of appearance as a plane color, and says that, like a surface color, "it spreads out in space in a bidimensional, plane-like manner" (9, p. 8). Martin, on the other hand, says that the "film" tends "more strongly toward tridimensionality than Katz' bidimensional *Fläche*" (12, p. 479). Bühler, like Katz, speaks of this color as a plane color (*Verdichtungsfläche*) (2, pp. 53-57), while Müller, like Martin, emphasizes the tridimensional character of the "original visual sensations." "Like the sensation of the black of the closed eyes, they have something more or less bulky about them" (13, p. 71). We believe that these differences might be resolved if one kept continually in mind the notion that the film color or any other mode of appearance should not be objectified, should not be made to assume the properties of an independent physical object. It is entirely possible that a color which appears out in front of us, behind a double screen, might from one attitude be seen as plane-like, and from another as somewhat tridimensional. And, similarly, on the side of report, an observer characterizing a reduced color with one standard in mind, say, a definitely bulky color, might speak of a film color as plane-like; while with another standard in mind, say a definite surface, he might speak of it as somewhat tridimensional. We believe it would be more accurate not to try to identify these two guises of the reduced color—its apparent tridimensionality and its apparent plane-likeness—but instead to realize that they may arise as a result of somewhat different observational attitudes, and should therefore be recognized as somewhat different experiences.

The Light-dark Integration. Certain features of the phenomena that we have reported under this head are not altogether new. They are new only in the fact that they are found to occur in luminous appearances, and in the peculiar form they take in these appearances. The form of light-dark integration in which a greater brightness is seen as lying further back in the visual object is evidently but a special case of the "appearance behind each other of colors in space," which has been reported and discussed many times before—by Helmholtz (5, p. 276), Schumann (15), Bühler, (2, pp. 40-45), Henning (6), and Fuchs (4)—to mention only a few. It has not, however, so far as we know, been previously reported in connection with luminous appearances. Moreover, as observed here, the phenomenon shows certain peculiarities. The colors which are seen behind each other in space are not apprehended as distinct and separate entities as, for example, when we see an object through a colored glass, or through a bulky liquid. Rather, the different qualities involved occur together as parts of a single unitary visual object. In some cases, however, there is a little greater spatial distinctness, as when a veil-like darkness is seen in front of a brighter mass, while in others the phenomenon consists merely in an increased brightness of the visual object as one penetrates further back into it. A phenomenon of similar nature was reported by Martin. Some of her observers, in describing the appearance of the "cloudy haze," which we have already referred to, reported differences in hue between the near and far portions of the visual object (12, p. 474). In connection with these integrations, it may be recalled that our observers spoke of the brightness behind as being a homogeneous, pure brightness, although a homogeneous stimulus could not have been present to the eye. In some cases, as in that of the heated copper sheet, this pure brightness seemed to cohere as a somewhat separate core or mass. There is probably involved, then, some factor like that which Bühler suggests in connection with the coherence plane. There, a number of points which are non-continuous with each other seem to lie in a plane; there is present a "coherence factor" of some sort which leads them to be apprehended as a particular kind of unified total. Here too, then, since the brightness behind is taken as existing homogeneous and pure further back in depth—in a sense, therefore, as a total somewhat discrete from the relatively darker quality in front—some sort of coherence factor may be involved in the perception. Light-dark integrations of the second type, those in which both qualities are distributed throughout the extent of the visual object, have also been reported by Bixby in the phenomenon of luster (1). We shall refer to his report in some detail later on.

V. Luminosity and its Varieties

We may now specify, in terms of the phenomenological variables which we have defined, the properties that all luminous appearances have in common, the different types of luminosity that may be differentiated, and how these types merge into each other by way of series of intermediates.

a) *Luminous Appearances.* Our reports show the common characteristics of luminous appearances to be: first, a brightness greater than that of the immediate surroundings, i.e., the background against which the color appears; and, secondly, a volumic character. Two things should be kept in mind with reference to our use of the term volumic. The volumic character that is possessed by our luminous appearances is not of the same sort as is possessed by a solid object or even by an opaque cloud. With solid objects and opaque clouds, which also appear to occupy an extent in space, one does not seem to be able to look into the depths of the experience; he is stopped by an opaque surface. But, as we have pointed out, there is no definite frontal limit to our visual object; one can penetrate its depths to some extent; one's gaze is not stopped off short by a hard impenetrable plane. It should also be realized that the term volumic does not refer to true bulky colors. It is significant that two of the observers, who saw the central green bulb of the Geissler tube as a bulky color, stated positively that it did not appear luminous.

F: (Geissler tube, central bulb) The green above and below the pink band was definite bulk in that I could see the back plane, see where it stopped. The green bulk was neither glow nor luminous. *H*: (Geissler tube, central bulb) The green outlying portions are pure bulk. They are just like solid green glass, or solid green ice. They show a fairly definite surface, with high lights and little points of brightness. But there's no luminousness to the green. It certainly is not glow.

Those appearances which, although brighter than their immediate surroundings, were not volumic in character, were not designated as luminous. We have already, in our exposition of the volumic type of appearance, listed the stimulus-objects for those appearances which possess the characteristics essential to luminosity; and, in our exposition of the non-voluminous appearances, we have named the stimulus-objects which do not possess the two necessary characteristics.

b) *Varieties of Luminous Appearance.* The appearances reported as luminous were of three types. We shall refer to them as α, β, and γ glows respectively. By *α glow* we refer to those luminous appearances which are of a high, homogeneous, continuous brightness, almost pure white in quality. Examples of this type were furnished by the tungsten filament with currents of 1 to 1.125 amperes through the lamp. Other good examples are automobile headlights or street lights observed from a distance of 50 to 200 yards. By β *glow* we refer to those light-dark integrations in which a greater brightness is localized within the depths of the visual object, e.g., the appearance of a red-hot iron. By γ *glow* we refer to light-dark integrations in which the lighter and darker qualities are interfused throughout the entire extent of the visual object, as in the appearances of the Aurora and Geissler vacuum tubes. A fourth possible class, which we merely mention now and shall discuss later, comprises the flames. It is not coordinate with the classes named above, and its members may also find a place in one of these classes.

The only type of luminosity which the observers agreed in designating as glow was that which we have called the β glow. We may therefore regard it as the experience referred to by the term glow in everyday life. The α glow was sometimes called incandescence, sometimes sparkle, and sometimes white glow. The γ glow usually was referred to as simply luminous. In the absence of any special common names for these last two varieties, we refer to them all as glows as indicated above. We find support for this terminology in the tendency of physicists to designate all luminous appearances other than flames as glows;[6] in the absence of any other sanction we follow the same procedure.

Between the α and β glows, we find a continuous series of intermediates. We provided these by starting with a current of 0.57 amperes through the tungsten lamp and increasing it by small steps in successive exposures. With the initial current (0.57 A.) the experience was that of a very good glow (β glow); the visual object was distinctly massive and volumic, and showed a greater brightness in the depths of the experience; it was described as a thin film partially veiling a greater brightness behind. But as the current was increased, and the brightness of the experience as a whole increased, this veil became more and more tenuous, but the greater brightness was still localized in the depths of experience. As this change progressed, the greater brightness behind also became whiter in quality, and the orangish red color characteristic of the less bright experience gradually disappeared. Finally, with the highest currents, there was present only a homogeneous mass of dense white brightness, the α glow. The β glow extends also in the opposite direction of brightness by way of a continuous series. The successive members of this series were provided by the heated iron rod as it decreased in temperature and faded out, and by the tungsten filament with successively smaller currents through the lamp. The iron rod illustrates particularly well the nature of the transition from the glowing to the non-glowing experience. At the beginning it is seen as β glow and is described as showing a greater brightness within the depths of the experience. At some point before it has completely faded out, it is apprehended as not glowing. Yet the observer is not able to detect a definite point at which it ceases to be glow and becomes not-glow; the change in the visual object seems to be continuous. When the point is reached, however, at which the observer considers the experience as not glowing, he also finds that the volumic character and the greater brightness in the depths of the experience have disappeared. There is thus marked out a series ranging from bright α glows at one extreme through the entire extent of the β glows to a non-volumic, non-glowing experience at the other extreme.

---

[6] The following references illustrate this tendency: (17, 20, 18, 8, 14).

Certain variations in hue were found to occur among the different types of luminous appearance. The objects showing the α glow were seen as white or yellowish white.[7] The β glows, except in the case of the phosphorescent substances, were of an orangish hue varying between reddish orange at the low and yellowish orange at the higher brightness.[8] The γ glows appeared in a number of varied hues. To determine whether luminous appearances of the β type would still be seen as "typical" glows when the hue was varied, we arranged to have our observers look at the tungsten filament through differently colored filters.[9] The resultant experiences were in every case designated as glows and always showed the characteristic phenomenon of a greater brightness in the depths of the visual object.

---

[7] This is explained by the fact that all light waves of high intensity arouse a white quality. As the wave-energy increases, all colors pass over into a white by way of either yellow or blue, depending on the spectral distribution of the stimulus.

[8] The presence of such hues is explained by the fact that heated solid substances were the stimulus-objects. Such substances, as they rise in temperature, begin to emit energy falling within the visible spectrum. At first, long waves (red) predominate, but soon the intensity of the shorter waves increases and the spectral distribution approximates that of so-called white light. At the same time, the absolute intensity of the whole distribution increases. The color of the experience changes correspondingly from reddish to yellowish orange and finally to white.

[9] The filters used were deep red, orange red, orange yellow, pure green, blue green, and violet, all of which are included in the set of Wratten monochromatic filters manufactured by the Eastman Kodak Company.

The hues exhibited by our γ glows were red and blue. It should be possible by the use of vacuum tubes containing different gases to obtain a still greater variety of colors.

Concerning the γ glow a further question may be raised. Our description coincides in several respects with that given of luster by Bixby. Are the two accounts to be therefore considered contradictory? Not at all. Bixby attempted to describe luster considered apart from the surface with which it usually occurs. His conclusion was that "luster is always independent of the surface of the lustrous object; in texture it stands somewhere between film and bulk; its quality is a peculiar commingling or sifting-together of dark and light" (1, p. 139). An examination of the reports of Bixby's observers shows that the last two phrases, if translated into our terminology, would mean that "luster" is somewhat volumic in character, and qualitatively is a light-dark integration in which both qualitative aspects are intersifted throughout the extent of the experience. These two characteristics are also found in our γ glow. But it is quite possible that when the phenomenon found by Bixby in luster does occur alone, and apart from any surface, it will be seen as luminous. This is verified by an examination of the reports of his observers. In many of them the "lustrous phenomenon" is reported as "luminous" or "glowy."

Flames. We have found that in their inner constitution, in the manner in which they appear to emit light, flames do not show any peculiar pattern; they resemble different types of the luminous appearances we have distinguished. The *blue flame,* for example, resembles the γ glow, and the *yellow flame* the intermediates between the α and the β glows. Why, then, are flames so often marked off as different from other forms of luminosity? Katz seemed to distinguish them from glows on the basis of sharpness of contours, but Bühler denied that the outlines of a glowing body are sharp. We, too, have found that this is not universally the case. Some of our glowing bodies did, it is true, show rather sharp outlines, as e.g., the heated rod in its first stages. But the contours of the glow of the copper sheet or of the iron bar in the later stages were, on the contrary, extremely indefinite and hazy, much more so in fact than those of the yellow candle-flame; the contours of the latter were very sharp and clean-cut.

Our suggestion for distinguishing flames from other luminous bodies would be, rather, on the following basis. Flames show a constant intermittent movement up and down and from side to side. This gives them an extremely light and unstable character compared with glows which show no such constant shifting of position and contours. The amount of this movement varies in different flames, being more pronounced, for example, in that of a Bunsen burner than in a candle-flame in a room free of air currents; but present to some extent in all flames. Two other marks which, though not universal, are quite general are the characteristic form, a tapering of the flame toward its outer end, and the presence of a relatively darker region near the source of the fuel supply. These characteristics, together with the peculiar movement, mark flames off from other luminous appearances and enable us to recognize them as flames.

VI. Brightness as a Condition of the Luminous Appearance

In the past much importance has been attached to the brightness of a visual object as a factor in the appearance of luminosity. It has not been maintained, however, that the relation is simple. There has been a uniform tendency to consider the impression as dependent, not upon an absolute degree of brightness of the experience but upon the brightness as related to the conditions of illumination. Hering, we will recall, stated that a *sufficient* condition for the appearance of luminosity was that an experience be seen as brighter than the purest white things under the same conditions of illumination. As we have already pointed out, this was taken over by Katz and Bühler as expressing a *necessary* condition for the impression.

A simple experiment shows, however, that this version of Hering's law is not correct. The anode end of the Aurora tube can be seen as luminous even though it is not as bright as a white thing under the same conditions of illumination. But the difference between the luminous and the non-luminous impressions, in this case as elsewhere, is that the luminous experience appears volumic, whereas the non-luminous experience, although brighter, appears as a flat surface.

Our arrangement was as follows: Directly behind the rectangular hole in the screen was placed the anode end of the tube, and a couple of inches to the right, on the screen, a piece of Baryta white paper (Stoelting 17). The observer sat at a distance of 135 cm. from the screen, directly opposite a point midway between the paper and the hole. A 40-watt frosted Mazda lamp at a distance of 2.45 m. from the screen and 65 cm. to the right of the observer supplied the illumination. The observers, F, H, and K, saw the Aurora tube as luminous and the white sheet of paper as not luminous, although the conditions of illumination were the same and the Aurora tube did not appear as bright as the white paper. The brightness of the Aurora tube was 0.38 millilamberts, and that of the white paper was 0.6 millilamberts, as measured with a Holophane Lightmeter.

It is curious that Katz and Bühler, who have granted a certain validity to their version of Hering's law, have nowhere cited evidence for it. Katz merely states that the law holds in most cases, but he does not state the difference between the cases in which it holds and those in which it does not (9, p. 28). Bühler, while failing to supply any evidence, does expound a theory of the conditions underlying all impressions of illumination, of which the impression of luminosity is one. We must, therefore, examine his theory insofar as it attempts to account for the impression of luminosity.

Bühler bases his theory on experiments of L. Weber on the scattering of light by what he calls the air-plankton. This is defined as the air, together with all the small particles it contains. A certain amount of the light passing through this air-plankton is scattered in all directions, and so we can think of the air-plankton as a source of light. Weber contends that light emitted in this way, by even relatively small layers of air, is photometrically measurable; and he devised a photometer for measuring the *albedo* of the air-plankton, that is, the ratio of the intensity of the diffuse light sent out by the *plankton* to the intensity of the incident light. He then made some measurements of the albedo of certain extents of the air-plankton under different weather conditions (19).

On the basis of these results, Bühler contends that scattered air-light is the stimulus for the impression of empty space and the underlying factor for the impression of illumination that may arise in any part of this space. We see a body, he contends, not only as of a certain brightness which corresponds to the intensity of the light reaching the eye from the body, but as weakly or strongly illuminated. That is, if a black object (i.e., a weakly reflecting object) and a white object (i.e., a strongly reflecting object) were capable of being differently illuminated so that the same amount of light reached the eye from both bodies, we should not see them as of equal brightness, but should see the black object as a strongly illuminated black, and the white object as a weakly illuminated white (2, pp. 74f). In other words, the albedo of the body surface can be apprehended by the observer. This is made possible, he says, by the air-radiation, which will vary directly with the intensity of the illumination and so furnish a measure of it. "Let us grant," writes Bühler, "that the body-radiation is not the only radiation reaching the eye at the time, but that the air before and around bodies also radiates, and, indeed, so as to be noticeable to the eye. Consider the intensity of this air-radiation as univocally determined, as an expression of the intensity of light prevailing in space, and thus also of the intensity of illumination striking the objects. . . . This effective air-radiation can, then, furnish the measure sought for. . . . What is not possible directly is attained indirectly; in a very simple and technically elegant way, an equivalent is furnished . . . . for the important quotient, $\frac{\text{the intensity of reflection}}{\text{the intensity of illumination}}$" (2, p. 97).

How does Bühler relate this theory to the impression of luminosity? In his reply to Katz' review he says: "There must be a luminosity limen, and, indeed, a sliding limen depending on the prevailing intensity of illumination. If we, in some way, determine the magnitude of the illumination factor in the impression, we thereby also determine a maximum value for the radiation emitted by reflection from the illuminated plane. To put it briefly, the plane cannot reflect more than it receives. But if its radiation exceeds this value, then the impression of luminosity arises" (3, p. 182f).

The last two sentences contain the gist of the doctrine. If the intensity of the light received from a body by the eye exceeds the amount striking the body, then the body is seen as luminous. And the amount striking the body, the intensity of illumination, is registered in the observer by way of the air-light, which serves as a measure of the illumination. This theory makes understandable Bühler's belief that a body would have to be seen as brighter than the brightest possible white thing to be seen as luminous. For the brightest possible white thing under a given condition of illumination would be the brightness of the body that came closest to reflecting all the light striking it. Any body brighter than this would have to be radiating light of its own in addition to reflecting that of the illumination. It would, therefore, be physically self-luminous and, according to Bühler, would be seen as luminous.

The theory seems to rest on several assumptions. First, that the air-light is superliminal for all conditions in which the impression of luminosity may occur; therefore, for even the thinnest layers of air. Secondly, that there is a neural device which permits the radiation from the air in front of a radiating or reflecting body to be separately effective in the observer from the radiation from the body that has traversed the entire layer of air. Thirdly, that there is a nervous mechanism which provides the impression of luminosity as soon as the radiation received by the eye from the body exceeds by a sufficient amount the radiation received from the air. For the intensity of the air-light is considerably less than the intensity of illumination which it represents. In order, therefore, for radiation from a body to be apprehended as coming from a luminous body, it must exceed, not merely the intensity of the light received from the air, but the intensity of the illumination which the air-light represents. The theory must assume, then, that there is some mechanism in the eye or the nervous system which, figuratively speaking, knows by *how much* the intensity of illumination exceeds the air-light received in the eye as a result of that illumination. But the mechanism must be still more clever. The air-light varies, not only with the illumination, but with the condition of the air—with the amount of moisture and dust that it contains. The nervous mechanism, therefore, which uses the air-light as an index of the illumination would have to make corresponding shifts in accordance with weather conditions. Finally, the theory assumes that the radiation from a body *must* exceed the intensity of illumination in order that it be seen as luminous, and that *if* the light sent out from the body does exceed the intensity of illumination, it will *necessarily* be seen as luminous.

Let us consider these implications of Bühler's theory, starting with the last. (*a*) Must the intensity of light radiated from a body exceed the intensity of illumination in order for the body to be seen as luminous? Evidently not, as shown by the experiment with the Aurora tube which we have already reported. The Aurora tube was less bright than the white paper. Therefore, its light radiation must have been less intense than that from the white paper. But all of the light from the white paper consisted of reflected illumination. Hence, the intensity of the light radiated by the tube must have been less than that of the illumination. Yet the Aurora tube was seen as luminous. (*b*) If the light-radiation from a body does exceed the intensity of the illumination, will the body necessarily be seen as luminous? Again, we can cite evidence supporting a negative answer. Our heated iron rod, shortly before it faded out completely, was seen as non-luminous, although still visible. The intensity of illumination was zero. The intensity of radiation from the rod, since the rod was visible, was obviously greater than zero, and hence, greater than that of the illumination. Katz, too, states that a physically self-luminous body in a dark room may appear as a visible film color, and yet not as a luminous color (9, p. 28).

Another simple experiment illustrates still more prettily the dependence of the appearance of luminosity upon a volumic character rather than upon conditions of illumination of the sort suggested by Bühler. We were able by a simple device to cause a stimulus-object, which had previously been seen as luminous, to cease to appear luminous, even though all conditions of radiation remained the same. The stimulus for this experiment was the heated copper sheet. It was seen as glowing and luminous when set up in the usual way behind the screen with a 16-cm.-diam. hole, which permitted the hazy contours of the visual object to be seen. When, however, we used in our screen a card with a hole 6 cm. in diameter, which permitted only a rather uniform central portion of this visual object to be seen, the observers' experience was not that of a volumic, glowing visual object, but that of a reddish-orange translucent plane stretched across the hole. It appeared like a colored glass through which light was coming from behind. The observer sat at a distance of 2.5 m. from the screen. The conditions of light radiation and illumination were the same in both cases, yet the impression of luminosity was present in one case and not in the other. It should have been present in both cases according to Bühler's theory, since the intensity of light radiated from the body exceeded the intensity of the illumination, which was zero. Although the observers were not instructed to do so, and did not consciously try to fixate at any given distance, it seems that the circumstances were such as to lead them to focus on the edge of the hole, and it was then that the experience occurred as we have described. One observer, H, reported that when he forced his fixation further back, beyond the screen, he saw the experience as massive, like a ball, which then was seen as glowing.[10]

We have now disposed of Bühler's factual statement. Let us consider briefly the difficulties with the other principles implied by his theory. For the existence of a physiological mechanism which would permit the air-light to serve as a measure of the illumination we know of no evidence. Nor is there evidence of the existence of a mechanism by means of which the air-light before an object would be separately effective in the eye of the observer from the light which proceeds directly from the stimulus-object. We should expect Bühler to suggest the nature of such necessary mechanisms. Finally, there is the question whether the air-light emitted by relatively small layers of air is psychologically effective at all, regardless of whether it can serve as a measure of the illumination. Katz claims that it is not; that is, he says that the quantities of light emitted by such air-planktons are sub-liminal. He points to experiments of his own in which he found that, under clear air conditions and with layers of air about ½ m. in thickness, the light of the air-layer was under the relative difference limen, and could be considered a physiologically irrelevant magnitude.[11]

Theories of the factors underlying a given perception should, we believe, begin with a complete description of the perception, rather than with a search for a parallel in experience to a supposed physical stimulus. Then one may proceed more surely by investigating the physical and physiological conditions underlying the different characteristics observed in that perception. One is less likely to be carried away by some special physical fact or to neglect the extreme importance of attitude as a factor in the perception. So, in the case of luminosity, we believe that an investigation of its conditions should proceed by determining the physiological and physical conditions underlying the volumic types of appearance and the different types of light-dark integration. In this way much fruitless theorizing might be avoided.

---

[10] The appearance of the reddish orange of the heated copper sheet as a flat plane where its contours are controlled in the way described furnishes an instance of what Bühler would call a coherence plane arising as a result of a frame-effect (*Rahmenwirkung*) (2, p. 62).

[11] See (10). The results of these experiments were published later (11).

VII. SUMMARY

Previous studies of luminous appearances have either been incidental to other investigations, or have not been concerned specifically with the phenomenological properties of these appearances. The present study was undertaken expressly to determine the phenomenological characters of a representative group of luminous appearances.

Our stimuli gave us a large number of luminous—along with a number of non-luminous—appearances. We found the luminous appearance to be characterized by (*a*) a volumic character of the visual object and (*b*) a brightness of the visual object greater than that of its immediate surroundings.

This volumic character differs both from bulky colors as defined by Katz and Bühler and from the solid impression of objects. Bulky colors, while transparent, have definite front and back limits. Solid objects seem to occupy a definite portion of tridimensional space even when opaque. The luminous appearance, however, is visually penetrable and seems to occupy an extent in depth without definite front or back limits. An incidental result of the comparison of this characteristic with those of the film color was the discovery that the description usually given of the reduced film color is not true of colors of high brightness. There is need, therefore, of an investigation of reduced colors of different quality.

With reference to Hering's law, that "we see something as luminous or lustrous when it is sensed as brighter than pure white things in the same illumination and when under the circumstances the greater brightness cannot be taken as incident light," we have shown (*a*) that this law was misinterpreted by Katz and Bühler, and (*b*) that their formulation of the law is not verified by experiment. We have also indicated certain difficulties in Bühler's psychophysical theory of the conditions of luminosity.

The luminous appearances observed are of three types. The first is characterized by homogeneous brightness (e.g., the appearance of a tungsten filament with a current of 1.125 amperes), the second by a light-dark integration with a greater brightness localized within the depths of the visual object (e.g., the appearance of red-hot iron), and the third by a light-dark integration with the light and dark interfused throughout the entire extent of the visual object (e.g., the appearance of the anode end of an Aurora vacuum tube). We have designated these types as α, β, and γ glows, respectively.

The light-dark integration presents a peculiar pattern. The light and dark are both present in the visual object but they cannot be stably localized at definite points; they merge into each other indefinitely and are not sharply separated by definite contours. We have here then something quite different from the familiar pattern of light and dark surfaces.

The β glow was the only one uniformly characterized as "glow" by our observers, and may be regarded as the experience indicated by this term in everyday life. A continuous series of β glows can be made out ranging from one that is hardly distinguishable from the homogeneous, highly bright α glow to one with a minimal internal brightness.

Except for the α glows, which are a pure white or a yellowish white of high brightness, the different types of glow may exhibit all possible variations in hue.

While flames have previously been marked off as a special kind of luminous appearance, we have found that they fall into our classification of glows. We have been unable to corroborate Katz' specification of flames as lacking the well outlined form of glowing bodies. The difference of flames from other luminous appearances consists, rather, in the peculiar movement exhibited by flames.

REFERENCES

1. Bixby, F. L. A phenomenological study of luster. *J. Gen. Psychol.,* 1927, **1**, 136-174.
2. Bühler, K. Die Erscheinungsweisen der Farben. Jena: G. Fisher, 1922. Pp. x + 211.
3. ————. Gegenbemerkungen. *Psychol. Forsch.,* 1924, **5**, 182-188.
4. Fuchs, W. Experimentelle Untersuchungen über das simultane Hintereinandersehen auf derselben Sehrichtung. *Zsch. f. Psychol.,* 1922, **91**, 145-235.
5. Helmholtz, H. Treatise on physiological optics. (Edited by J. P. C. Southall.) (Trans. from the 3rd German edition.) (Volumes 1 and 2.) Menasha, Wis.: Optical Soc. Amer., 1924. Pp. xxi+ 482; viii + 480.
6. Henning, H. Ein optisches Hintereinander und Ineinander. *Zsch. f. Psychol.,* 1921, **86**, 144-174.
7. Hering, E. Der Raumsinn und die Bewegungen des Auges. Hermanns Handbuch der Physiol., 1879, **3**, Teil 1, 343-601.
8. Houstoun, R. A. A treatise on light. London: Longmans, Green, 1924. Pp. 307.
9. Katz, D. Die Erscheinungsweisen der Farben und ihre Beeinflussung durch die individuelle Erfahrung. (*Zsch. f. Psychol.,* Ergbd. 7.) Leipzig: Barth, 1911. Pp. xviii + 425.

10. ————. Referat (K. Bühler's Handbuch der Psychologie—Die Erscheinungsweisen der Farben). *Psychol. Forsch.*, 1924, **5**, 172.

11. ————. Neue Beiträge zu den Erscheinungsweisen der Farben. *Zsch. f. Psychol.*, 1924, **95**, 129-136.

12. Martin, M. F. Film, surface, and bulky colors and their intermediates. *Amer. J. Psychol.*, 1922, **33**, 451-480.

13. Müller, G. E. Abriss der Psychologie. Göttingen: Vandenhoeck & Ruprecht, 1924. Pp. iv + 124.

14. Nichols, E. L., & Howes, H. L. The blue glow. *J. Opt. Soc. Amer.*, 1922, **6**, 42.

15. Schumann, F. Dimensionen des Sehraumes. *Zsch. f. Psychol.*, 1921, **86**, 258-277.

16. Stumpf, C. Die Attribute der Gesichtsempfindungen. *Abh. d. Berlin. Akad.*, 1917, *Phil.-Hist. Kl.*, No. 8.

17. Thomson, J. J. Conduction of electricity through gases. Cambridge, 1903. (Pp. 433, 443.)

18. Watson, W. A. A textbook of practical physics. London: Longmans, Green, 1022. (P. 307.)

19. Weber, L. Die Albedo der Luftplankton. *Ann. d. Physik*, 1916, **4**, 427-449.

20. Wood, R. W. Physical optics. New York, 1911. (P. 564.)

Cornell University
Ithaca, New York

LES APPARENCES LUMINEUSES

(Résumé)

Les apparences lumineuses montrent en général (1) un caractère de volume et (2) une clarté de l'objet visuel plus grande que celle de ses environs les plus proches. Le caractère de volume diffère des couleurs volumineuses comme définies par Katz et Bühler, et de l'impression solide des objets.

On montre que la loi Hering des conditions de l'apparence lumineuse a été mal interprétée par Katz et Bühler, et que leur nouvelle forme de la loi n'est pas vérifiée par l'expérience. On indique certaines difficultés dans la théorie psychophysique de Bühler des conditions de la luminosité.

Les apparences lumineuses observées sont de trois types. Le premier est caractérisé d'une clarté homogène, le deuxième, d'une intégration claire-obscure avec une plus grande clarté localisée au fond de l'objet visuel, et le troisième, d'une intégration claire-obscure avec les deux qualités distribuées sur tout l'objet visuel. Ces types sont appelés les lumières ardentes α, β, et γ respectivement. La lumière β peut être considérée comme l'expérience type ou bien connue de la lumière ardente comme elle se montre dans la vie quotidienne, par exemple, l'apparence d'un fer tout rouge.

Les flammes tombent dans la classification des lumières ardentes, et l'assertion de Katz que les flammes n'ont pas la forme bien dessinée des objets ardents ne se trouve pas corroborée. La différence entre les flammes et les autres apparences lumineuses consiste plutôt en le mouvement particulier montré par les flammes.

Kreezer

DIE LEUCHTENDEN ERSCHEINUNGEN

(Referat)

Die Merkmale leuchtender Erscheinungen sind: 1.) ein voluminöser Charakter und 2.) eine Helligkeit des Sehobjektes, die grösser ist als die der unmittelbaren Umgebung. Der voluminöse Charakter ist von dem dicker (bulky) Farben, wie sie von Katz und Bühler definiert wurden, und von dem Eindruck fester Körper verschieden.

Man zeigt, dass das Heringsche Gesetz über die Bedingungen von leuchtenden Erscheinungen von Katz und Bühler falsch interpretiert wurde, und dass ihre Wiederformulierung des Gesetzes durch das Experiment nicht verifiziert wird. Es werden gewisse Schwierigkeiten in Bühlers psychophysiologischer Theorie über die Bedingungen der Helligkeit hervorgehoben.

Die beobachteten leuchtenden Erscheinungen gehören drei Typen an. Der erste lässt sich als homogene Helligkeit, der zweite als Hell-dunkel-Integration mit grösserer Helligkeit innerhalb der Tiefen des Sehobjektes, und der dritte als Hell-dunkel-Integration mit Ausbreitung der beiden Qualitäten über die Gesammtausdehnung des Sehobjektes charakterisieren. Diese Typen werden als α, β, γ Glut bezeichnet. Man kann die β Glut als die typische oder bekannte Erfahrung der Glut betrachten, so wie sie im alltäglichen Leben vorkommt, beispielsweise die Erscheinung des rotglühenden Eisens.

Flammen fallen in die Klassifikation der Gluten, und die Katzsche Erklärung, dass Flammen die gutumrissene Form von glühenden Körpern nicht haben, bestätigt sich nicht. Der Unterschied zwischen Flammen und anderen leuchtenden Erscheinungen besteht eher in der den Flammen eigentümlichen Bewegungen.

Kreezer

A QUANTITATIVE STUDY OF BINOCULAR COLOR VISION*

From the Psychological Laboratories of Clark University

Dorothea E. Johannsen

I. Introduction

A survey of the studies upon the subject of the binocular perception of colors reveals a lack of uniformity both in the results obtained and in the interpretation of these results. Experientially there are three fundamental ways in which a color observed with one eye can affect another color simultaneously observed with the other eye. The terms commonly used to denote these three effects are dominance, rivalry, and mixture, but there is a divergence of opinion as to the conditions under which each occurs. Dominance is the suppression of one field by the other; usually the less dominant color appears in small and rapidly shifting patches. Mixture is the appearance of a color lying approximately midway between the two component colors. Rivalry is usually considered to be the more or less rapid alternation between dominance of one color and dominance of the other, but rivalry between mixture and dominance occasionally occurs.

The investigators of these problems may be divided into three groups according to the interpretation which they put upon experimental results. One group finds that the binocular observation of colors always produces an intermediate color; another group finds that mixture never occurs, but only rivalry; while the third group finds that dominance, mixture, or rivalry may occur, depending on physical and physiological conditions. In general, the older experimentalists (Haldat, 16, and Dove, 10) reported that mixture could always be obtained. Helmholtz (17, Vol. III, pp. 516-527) found only rivalry, and for him and the group under his influence the whole question of rivalry is very largely bound up with the problem of attention. The more recent studies tend to show that the

---

*Accepted for publication by John Paul Nafe of the Editorial Board and received in the Editorial Office, December 31, 1929.

phenomena observed are dependent upon the conditions of stimulation; Dawson (7) summarizes the work done up to 1911.

II. Formulation of the Present Problem

The present problem is limited primarily to a quantitative study of the physical conditions which determine the phenomena occurring when different colors are observed with the two eyes simultaneously. Previous investigations have shown that hue and intensity are among the most important factors in determining whether mixture or rivalry results; therefore, these were the two factors which the apparatus was particularly designed to control.

III. Apparatus

The investigation was limited to a study of transmitted color; contours were eliminated save for the boundaries of the observed fields. Transmitted instead of reflected light was used because of the greater opportunity it presents exactly to define the amount and length of the light waves which are permitted to impinge upon the eye. In view of the difficulty of eliminating all white light from spectral colors, gelatine filters were used, although none that are perfectly monochromatic are obtainable. An added advantage of the filter light is the very much wider variation in intensities which it permits. The Wratten monochromatic light filters seemed to present the best opportunity for obtaining a light which could be controlled both as to hue and as to intensity. Filters 70 (red), 73 (yellow), 74 (yellow-green), 75 (blue-green), and 76 (blue), giving a fairly wide range of hues, were used. The two sets necessary to the problem were calibrated by the Eastman Kodak Company, in order to make sure that both members of each pair transmitted the same kind and amount of light; any difference in the transmission value of the members of a pair is less than the least perceptible photometric difference of 1.7%. Such transmission values are valid only when the source of the light with which the filters are used emits the particular quality of spectral radiation against which they are balanced. The light sources in this experiment were two 200-watt concentrated filament tungsten lamps having daylight blue glass, standardized by the Eastman Kodak Company, and were the same lamps that were used for calibrating the transmission values of the color filters. Lamp A was of 101.8, and Lamp B of 96.5 candle-power; the two balanced in color when Lamp A was operated at 115 volts and Lamp B at 112.8 volts.

Since the total transmission value of the filters varied, the brightness difference thereby produced was corrected by adjusting the distances from the lights to the diffusion screens; the brightness at the screen was calculated on the inverse square law. Three intensities of stimulating light were used, giving brightnesses of 54.2, 145.6, and 390.4 millilamberts at the diffusion screen, when there was no color filter in the apparatus. These brightnesses are in the ratio of 1:2.68:7.19, the central value being the geometric mean of the other two.

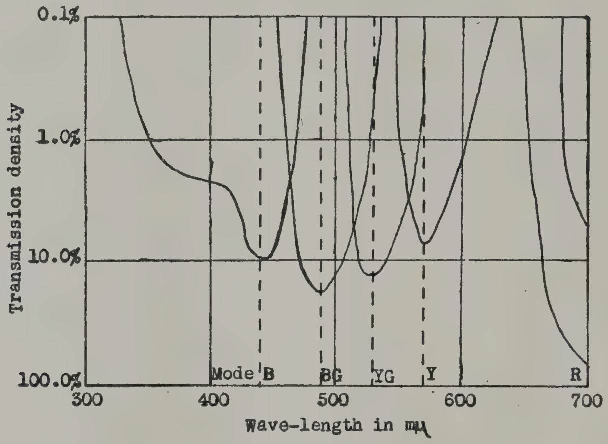


FIGURE 1

The filters vary from each other in another way: no filter is absolutely monochromatic, and the range of wave-lengths which each filter transmits is not the same, nor does it vary directly in proportion to the total transmission value. Only one filter (*Y*) permits non-consecutive groups of wave-lengths to pass. (Figure 1.) The filters are very much more adequate than spectral light-producing instruments in cutting out white light; practically no white light is transmitted.

There is a certain amount of overlapping in the range of wave-lengths transmitted by adjacent filters; this overlapping effect occurs only in those filters whose modes lie fairly close to one another, i.e., *B* and *BG, Y* and *YG,* etc. The wave-length of which the most is transmitted by any filter is called the *mode,* and this wave-length is the one which is used in all determinations involving the absolute wave-length of the color filters.

The fusion of the two fields into one binocular field was accomplished by using stereoscopic lenticular prisms. The filters were held directly behind these prisms in wooden frames which fitted so tightly against the front of the apparatus in which the prisms were set that no light could leak around the filter to the eye. The box containing the stimulating lights was divided into two compartments by a wooden partition in the median plane; the whole box was so constructed that light could neither leak into the outside room, nor penetrate into the compartment containing the other light. The lights were mounted on riders which were attached to a metric scale; they could thus be adjusted at any desired distance from the eye.

Screens made of perfectly clear and very thin glass (transmission value of 99%) were used to hold small black dots which served as fixation points. These dots were in focus when there was no color filter in the apparatus.

Diffusion screens of ground glass, whose transmission values were so chosen as to equate for the difference in the candle-power of the two lamps, were placed between the lamps and the color filters; when the candle-power of each lamp is multiplied by the transmission value of its corresponding diffusion screen the resulting figures stand in the ratio of 78.9 to 78.6. The rectangular field (6.1 by 6.6 cm.) was subtended by an optical angle of 16°20′ horizontally and 17°40′ vertically.

O reported by depressing finger-keys and a record of these reports was made on a smoked kymograph drum by means of writing-levers in circuit with the keys. Each trial took 58 seconds and covered a total distance of 50 cm. on the record.

IV. Observers

The observers were M. N. Crook (C), C. H. Graham (G), and R. W. Leeper (L), all graduate students in the Department of Psychology; all had previously observed in one or more experiments on visual phenomena, besides having had the regular laboratory course in experimental psychology. In some series Dr. J. P. Nafe

(N), Professor of Psychology, under whose direction this problem was done, also observed.[1]

V. PROCEDURE

The experiment was performed in a condition of light adaptation. The illumination in the observation room was diffuse artificial light (50-watt bulb suspended from the center of the ceiling). All other light was excluded. The observer sat in this room approximately 5 minutes before the observation was started.

Each *O* was given at least 10 hours of preliminary observation. The three reasons for the long training period were: to eliminate practice effect, to permit the drawing up of instructions suitable for the conditions, and to permit *O* to learn to know the experience and set up a criterion for distinguishing between the three fundamental modes of experience.

Every color was paired with every other color in such a way that in one set of observations the longer wave-length, and in the other the shorter wave-length fell on the right eye, making 20 observed pairs of hues. Where the intensities of the two observed colors were not the same, all possible combinations of hues were observed with the brighter light first on the right, and then on the left (40 observed combinations).

The instructions were read before each of the first three observations; they were somewhat detailed as to method of reporting and definition of what was to be considered as mixture and dominance. Before reporting on a new pair of colors *O* observed it for 58 seconds in order to determine what the colors were and to see what changes took place.

At *E's* "all right" signal *O* took his place at the hood, resting his chin and forehead against supports provided for them. At the "ready-now" signal the two lights were turned on simultaneously, and *O* reported the fluctuations, according to the instructions, by means of the finger-keys. Each pair of colors was observed 5 times in succession and the "results" are the averages of these 5 trials. At least 1 minute was allowed for recovery between each trial, and usually more time elapsed, since the voltage of each light had to be tested, the kymograph adjusted, etc. The kymograph record was measured with a cartometer. The distance over which one condition lasted during the 58-second run is recorded in centimeters. The conditions are classified according to the instructions and are as follows; (*a*) the dominance of one color, (*b*) the dominance of the other color, (*c*) mixture, (*d*) superposition of either color, and (*e*) a condition which could not be definitely reported in any of the other categories. The number of fluctuations was also counted.

---

[1] The writer is exceedingly grateful for the care and patience with which many hours of tedious observing were done.

VI. RESULTS

A. Effect of Various Aspects of the Stimulus upon the Phenomena Reported

The 20 possible combinations of the 5 colors may be divided into two groups of 10 each. The arrangements in which the longer wave-length falls on the right eye constitute the α series; those in which it falls on the left eye are called the β series. Since each set of arrangements consists of the same pairs of colors, reversed merely as to the eye which they stimulate, each pair was given an arbitrary number from 1 to 10. (Table 1.)

1. *Effect of Wave-Length*

a. Effect of wave-length as shown by dominance difference. Analysis of the data shows that the *hue* of the filters making up a pair has an influence on what will be reported. In general, the longer wave-length will dominate the combined visual field more than the shorter. As a measure of this effect, the total distance in centimeters in a given record over which the shorter wave-length dominated was subtracted from the distance over which the longer

TABLE 1

PAIRS OF COLOR FILTERS

Pair No.	Colors composing pair	Diff. in mμ	Diff. in j.n.d.
1	*Y* and *YG*	40	16
2	*YG* and *BG*	40	23
3	*BG* and *B*	50	27
4	*Y* and *BG*	80	39
5	*R* and *Y*	130	48
6	*YG* and *B*	90	50
7	*R* and *YG*	170	64
8	*Y* and *B*	130	66
9	*R* and *BG*	210	87
10	*R* and *B*	260	114

TABLE 2

EFFECT OF WAVE-LENGTH

Pair	Observer C			Observer G			Observer L			Observer N		
No.	D.D.	Mix.	Fluc.	D.D.	Mix.	Fluc.	D.D.	Mix.	Fluc.	D.D.	Mix.	Fluc.
					ALPHA SERIES							
1	....	48.5	1.0	....	49.0	1.0	....	49.0	1.0	....	49.0	1.0
2	9.8	37.5	3.2	.7	47.8	2.0	11.3	.4	9.8	....	48.2	1.0
3	....	49.0	1.0	....	49.0	1.0	....	48.4	1.0	....	49.0	1.0
4	46.7	....	7.0	13.5	33.6	4.6	27.4	.3	16.6	28.5	....	18.8
5	–2.9	14.3	10.4	....	49.0	1.0	9.9	....	32.0	26.3	20.3	21.6
6	21.5	27.0	2.4	16.6	26.9	10.6	24.7	16.4	17.4	45.4	....	4.8
7	27.0	....	32.0	–3.6	36.8	12.0	8.7	....	28.2	–9.8	....	23.6
8	46.7	....	6.2	35.7	....	16.8	26.4	....	23.4	40.8	....	9.4
9	25.5	....	42.6	18.7	....	14.4	16.9	....	33.4	4.5	....	23.4
10	27.2	....	41.4	13.0	....	21.6	29.5	....	33.0	13.5	....	24.2
					BETA SERIES							
1	....	47.8	1.0	....	49.0	1.0	....	49.0	1.0	....	49.0	1.0
2	....	38.5	1.0	4.7	38.4	7.2	.8	41.0	6.8	20.2	28.8	9.2
3	....	48.2	1.0	....	49.0	1.0	....	49.0	1.0	....	49.0	1.0
4	8.1	....	22.0	43.2	....	9.8	–2.8	....	13.0	13.3	....	19.2
5	–27.4	16.8	3.8	–5.9	40.5	5.6	12.7	....	11.8	11.2	19.4	13.8
6	12.7	11.7	11.4	18.8	24.6	8.8	3.1	....	25.4	24.2	....	19.2
7	6.7	....	38.6	–10.1	34.2	7.0	–9.7	....	24.6	–8.1	....	22.4
8	27.8	....	32.8	38.2	....	17.0	–5.5	....	28.2	17.5	....	20.0
9	19.5	....	36.4	–11.8	....	12.4	1.4	....	30.6	.3	....	20.8
10	21.2	....	47.0	22.2	....	20.0	1.3	....	38.2	1.8	....	26.6

D.D.—Dominance Difference
Mix.—Mixture
Fluc.—Fluctuations

wave-length dominated; this figure has been called the *dominance difference*. Table 2 gives the results for four *O*'s for both the α and β series. The column marked "D.D." gives the total distance over which the longer wave-length dominated the binocular field *in excess* of the shorter wave-length, i.e., the dominance difference.

The factor which suggests itself most strongly as a possible determiner of the size of the dominance difference which occurs with any given pair of colors is the amount of difference between the members of the pair. We may get an expression of this difference by measuring either the number of wave-lengths between the two, or the number of j.n.d.'s The amount of dominance difference, therefore, was correlated with each of these variables. (Table 3.)

Jones' (21) data for the number of j.n.d.'s in the spectrum were used; he nowhere states what the intensities of his stimulating lights

TABLE 3

CORRELATIONS

	Alpha Series								Beta Series							
	Wave-length				j.n.d.				Wave-length				j.n.d.			
Observer	C	G	L	N	C	G	L	N	C	G	L	N	C	G	L	N
D. D. orig.	.208	.030	.101	—.088	.386	.357	.333	.155	.684	—.315	.054	—.756	.405	—.310	.405	—.523
Mix. orig.	.800	.664	.867	.770	.806	.833	.818	.839	.818	.658	.806	.770	.830	.782	.809	.839
Fluc. orig.	.873	.770	.939	.794	.785	.891	.882	.721	.915	.635	.891	.818	.915	.786	.961	.833
D. D. focus	.304	—.530	——	——	.500	—.214	——	——	.185	—.101	——	——	.452	.148	——	——
Mix. focus	.691	.481	——	——	.709	.664	——	——	.800	.624	——	——	.806	.730	——	——
D. D. metronome	.018*	—.540	.541	——	.238	—.381	.309	——	—.172	—.564	.268	——	.095	—.262	.000	——
Mix. metronome	.876	.433	.858	——	.921	.591	.882	——	.800	.536	.876	——	.842	.745	.882	——
Fluc. metronome	.933	.485	.927	——	.858	.676	.515	——	.867	.385	.679	——	.858	.530	.770	——
D. D. no blink			—.125				—.143				—.101				—.429	
Mix. no blink			.836				.797				.921				.891	
Fluc. no blink			.945				.882				.889				.918	

*See text for added controls.

were. Smith (39), working with the purples, found that with the higher intensities more colors were discriminated. Since three different intensities were used in this problem it is probable that a different number of j.n.d.'s exists between the members of the same pair when the intensity is changed. Since the correlation between the wave-length difference and the j.n.d. difference between the members of the pairs used is .96, however, we seem justified in assuming that the number reported by Jones is accurate enough for the present purpose.

Correlations were computed by the rank-difference technique since there were only 10 cases; because of the difference between the α and β series they were correlated separately. The correlations are reported as ρ, an expression comparable to the r obtained by the Pearson product-moment method, though less reliable; the actual figures express a slightly higher degree of relationship than an r of the same size would indicate.

The correlation between the dominance difference and the difference between the members of a pair is not marked. For the α series, only one of the correlations between the wave-length difference and the dominance difference is above .20. The correlations between the number of j.n.d.'s and the dominance difference is higher, ranging from .155 to .386, only one being below .33. We have here an indication that there is a tendency for the longer wave-length in a pair, the members of which are rather widely separated, to dominate more than in a pair where the difference between the two colors is less.

In the β series we find much wider variations. Correlations between wave-length and dominance differences range between —.756 and +.684, for different *O*'s. The lack of consistency between *O*'s is puzzling; with such tremendous variations between *O*'s, it seems hardly likely that we have struck upon the most essential factor in determining the amount of dominance difference, although the variations in the correlations between j.n.d. and dominance difference are slightly smaller.

b. Characteristics of dominance differences

1) *Right- and left-eye effects.* The number of positive differences is larger in the α than in the β series; three possible explanations of this suggest themselves: first, a right-eye dominance, present in all *O*'s, and reinforcing the long wave-length dominance

on that side; secondly, a difference in the two sides of the apparatus; and, thirdly, a difference in the reaction-time of the two hands.

The last point is most readily disproved. In the first place, the dominance differences are too gross to be accounted for by such slight variations as reaction-times could supply. Reaction-time differences are measured in sigma; the dominance differences vary from 1 cm. to 48.0 cm. (i.e., from .116 sec. to 55.68 sec.). In the second place, the *O*'s used different techniques in reporting; L consistently reported the colors seen with the corresponding hands, while C consistently reported the shorter wave-length with the right hand. Both are right-handed; if the reaction of this hand is a trifle speedier, then we might expect that L's faster reaction with that hand would increase the positive dominance difference in the α series. In C's case, however, we would expect a *decrease* in the size of the dominance difference; his differences, however, are very large. This seems adequate proof that the difference is not dependent on the relative reaction-times of the two hands.

In order to check the second factor suggested, two control experiments were introduced.

a) *Control of diffusion screens.* The possibility of a selectiveness of the rays of light transmitted by the diffusion screens made it desirable to run a series of observations with them transposed. Were the diffusion screens differential as regards the light waves transmitted, the character of the light falling on the right color filter would be different from that falling on the left, and consequently a different type of stimulation might very well result from the same pair of filters when transposed from one side to the other. With these considerations in mind, the right diffusion screen and its corresponding light were shifted to the left side, and vice versa. Certain pairs showing a marked difference between the α and β series were repeated under these conditions. Most of the repeated pairs showed no reliable change in dominance difference; about half the shifts in the direction of the difference were toward greater positive differences for α and β series, and about half were in the opposite direction. This seems adequate evidence that differential transmission of the diffusion screens did not play a rôle in the differences between the α and β series.

b) *Control of stereoscopic prisms.* The stereoscopic prisms used were removed from an ordinary hand stereoscope, and it was possible that the apparent eye-difference was due to an inequality in the power of the two prisms; the prisms were, therefore, reversed with respect to the two eyes. A series of observations with the prisms reversed was run on each of two *O*'s. The differences obtained between this series and the original series are neither large nor consistent in direction. From this we are forced to conclude that any differences between the two prisms are too slight to lead to any consistent dominance difference.

The first suggestion for explaining the right- and left-eye differences apparently must stand. The question of a dominant eyedness, possibly dependent upon, or correlated with, dominant handedness, offers some interesting possibilities. It has been frequently suggested that there is a dominant eye which corresponds to the dominant hand (Downey, 12); all three *O*'s in this experiment are right-handed. If we now assume that the longer wave-length is the color which tends to dominate, but that the right eye is the dominating eye, then we have a situation where the two dominant forces may act to supplement each other (α series) or to counteract each other (β series). From the fact that even in the β series the long wave-length dominates in the majority of cases, we may conclude that the effect of hue is stronger as a determiner of dominance than is the eye.

c. Effect of wave-length as shown by mixture. The question as to whether or not the binocular observation of two colors will produce a mixture similar to that produced uniocularly by mixing the colors on a color-wheel has caused more controversy than any other one point. For the present problem mixture was defined as the occurrence of an "intermediate blend, which lies somewhere between the two colors and possesses the uniformity of texture and the stability of a single color."

The amount of mixture which was reported is also found to be a function of the hues involved. In the second column of Table 2 (headed "Mix.") the amount of mixture is reported. The figures are simply the total number of centimeters over which mixture was reported for the given pair.

Mixture tends to result from those pairs of colors which are closest to each other, whether the distance between them be measured in j.n.d.'s or in mμ. Certain pairs, *R* and *Y, B* and *YG, R* and *YG,* mix part of the time; these are the pairs upon which it is most difficult to report, for the mixture develops so gradually that it is difficult to observe its beginning. This ambiguity appears in the records as a greater variability in the judgments of mixture for these pairs. Table 3 gives the correlations for each *O*, for both α and β series, between the wave-length or j.n.d. difference and the amount of mixture reported for each pair. All these correlations are positive and range between .65 and .86; this is true for both the α and β series. We may, therefore, be justified in concluding that only those pairs the colors of which show a certain amount of proximity will mix.

From Figure 1 it will be seen that each filter does not transmit only one wave-length, but a group of wave-lengths (see p. 284). Pair 2 (*BG* and *YG*) transmits the largest percentage of identical wave-lengths, while Pair 1 (*Y* and *YG*) and Pair 3 (*B* and *BG*) transmit only slightly less. These are the three pairs of colors which show the largest amount of mixture. *R* and *Y* do not overlap in the same sense that the other pairs do, but the *Y* transmits a secondary group of wave-lengths in the *R* end of the spectrum (from 680 mμ to 700 mμ). This peculiarity of the *Y* filter may explain its inconsistent results when paired with the *R*; it may be that, for some *O*'s and under certain conditions of arrangement, intensity differences, etc., this amount of overlapping is sufficient to produce a mixture, while for others it is not. This pair of colors showed the greatest variety of reports and the most inconsistency in the amount of mixture, both from *O* to *O*, and from time to time with the same *O*, of any of the pairs.

Pair 6 (*B* and *YG*) showed a tendency to mix part of the time, particularly with the lower intensities. (See Tables 2 and 4.) The members of this pair, separated at the mode by only 90 mμ, do not overlap. With lower intensities discrimination is poorer; this may explain the apparent inconsistency of having colors without identical wave-lengths mix. Another explanation will be offered later.

d. Effect of wave-length as shown by fluctuations. The number of times the color fluctuated during each trial was counted and, as with the rest of the data, the average of five consecutive runs for each pair of colors recorded. The third columns of Table 2 show these results.

In Table 3 the correlations between the number of fluctuations

TABLE 4

EFFECT OF DECREASING THE INTENSITY ON BOTH EYES, AS SHOWN BY THE SIZE OF THE DOMINANCE DIFFERENCE AND THE AMOUNT OF MIXTURE

Observer L

		I_2 vs. I_2			I_3 vs I_3	
Pair	D.D.	Mix.	Fluc.	D.D.	Mix.	Fluc.
			ALPHA SERIES			
1	....	48.3	1.0	....	48.5	1.0
2	....	48.7	1.0	....	49.0	1.0
3	....	49.0	1.0	....	49.0	1.0
4	1.0	....	5.8	3.8	....	6.6
5	5.7	26.1	14.2	20.9	....	24.4
6	....	46.0	7.4	....	49.0	5.0
7	16.0	....	25.2	21.0	10.0	26.6
8	7.5	....	11.6	–1.8	....	13.2
9	–4.1	....	26.8	3.8	....	32.8
10	7.6	....	31.0	11.1	....	34.4
			BETA SERIES			
1	....	48.0	1.0	....	49.0	1.0
2	....	47.8	1.0	....	49.0	1.0
3	....	48.5	1.0	....	48.9	1.0
4	–14.4	27.6	2.0	–1.4	40.0	5.8
5	9.4	....	6.2	....	43.0	2.0
6	–8.4	16.6	15.6	–12.1	....	21.2
7	–22.8	.4	11.8	–8.9	....	8.6
8	–8.6	....	16.5	–8.8	....	25.0
9	–1.8	16.0	3.6	–12.8	....	25.4
10	–4.9	....	27.8	19.4	5.7	30.0

and the distance between the members of a pair are reported for each *O,* for both α and β series. The range here is between .63 and .96, but there are only two which are below .77. From this we may conclude that the number of fluctuations which occur in a definite unit of time is a function of the difference between the two colors, whether this difference is measured in terms of mμ or j.n.d.'s

2. *Effect of Intensity*

a. Effect of intensity as shown by dominance difference. As has been stated earlier, the intensities which were used in this experiment were in the ratio of 1: 2.68: 7.19, i.e., the first intensity was to the second as the second was to the third. The absolute brightness at the diffusion screen for the highest intensity was 390.4 millilamberts, for the second 145.6, and for the third, 54.2.

Two *O*'s (L and G) observed in the series where Intensity 1 (the highest) was paired with Intensity 3 (the lowest). Every pair in both α and β series was observed under two conditions, i.e., with the higher intensity on the right and the lower on the left, and the reverse. C observed in two series, Intensity 1 against Intensity 2, and Intensity 2 against Intensity 3.

Two different comparisons can be made to show the effect that intensity has on the reported phenomena. In the first place, the size of the dominance difference found when the higher intensity falls on the right side may be compared with the same figures for the same series of intensities reversed only with respect to the eye. In the second place, the size of the dominance difference found in the series in which the intensity ratio is 1: 1 may be compared with those in which it is 1: 2.68 (or 2.68: 7.19) or 1: 7.19.

Analysis of the reports of the simultaneous observation of Intensity 1 and Intensity 3 in the α series shows that for G in 6 out of 8 cases the dominance difference is larger when the higher intensity falls on the *right* eye; for L in 6 out of 7 cases the same thing is true. In the β series, however, for G in 5 out of 8, and for L in 7 out of 7 cases the larger difference occurs when the higher intensity falls on the *left* eye. This indicates that the higher intensity is a dominating factor which can act either to supplement or to counteract the effect of the long wave-length dominance, though it is not so powerful a factor as the difference in wave-length.

When this comparison is made for C's results for Intensities 1 and 2 and for 2 and 3, it is found that out of 16 cases in the α series 8 differences are larger when the brighter light falls on the right eye and 8 are larger when it falls on the left. Of the 15 comparisons in the β series 10 are larger when the brighter light falls on the left eye and 5 when it falls on the right. These slight differences indicate that these intensities are not different enough to cause any marked change in the amount of dominance in the two series.

A comparison of dominance differences reported where the intensities on the two eyes are the same with those where the intensities are different shows that in the former case the dominance difference is definitely larger than where the conflicting factor of different intensities is added (second comparison suggested to show the effect of intensity). Again, the dominance differences found in the series where the ratio of the intensities is 1: 2.68 (or 2.68: 7.19), as compared to those series where the ratio is 1: 1, are neither consistent nor large—a result which confirms the previous conclusion that this intensity difference is not large enough to produce a marked effect on the reported phenomena.

b. Effect of intensity as shown by mixture. The amount of mixture reported also offers an index of the effect of intensity. Analysis of the data shows that where the intensities are alike on the two eyes there tends to be a greater amount of mixture with the lower intensities for both α and β series; not one of the four exceptions to this statement is large and two are practically negligible.

If we now compare the amount of mixture occurring when Intensity 1 falls on one eye, and Intensity 2 on the other, with the amount occurring when Intensity 2 falls on one eye and Intensity 3 on the other (i.e., two series where the absolute intensities are not alike, but where the ratio between the two is the same), we find that the greater amount of mixture occurs in the series of lower absolute intensity.

A comparison of the results obtained from the simultaneous observation of Intensities 1 and 1, 1 and 2, and 1 and 3 is difficult to make because it necessitates in some cases comparing the results from different *O*'s, which is an unreliable proceeding, the individual differences in the amount of mixture reported being large. Insofar as we can generalize from such comparisons, we find that there tends to be more mixture when the intensities are alike than where they are of unlike, even though lower, absolute intensity.

In summary we may say that the less the absolute brightness the greater the amount of mixture, and the less the difference in intensity between the members of a pair, the greater the amount of mixture; and the latter factor is, perhaps, the more effective.

3. *Effect of Saturation*

a. The effect of saturation upon the observed phenomena. Another variable aspect of color is saturation or chroma. It has already been pointed out that the filters do not transmit the same number of wave-lengths, nor exactly the same total percentage of the incident light. The brightness difference which this produced was corrected by varying the distance of the light source from the filter according to the inverse square law. This did not, however, compensate for the fact that some transmit more different light waves (more nearly white light) than others; the former colors are, therefore, less saturated.

There are two aspects of the light transmitted by each filter which must be considered in making a measurement of its saturation. The first is the total number of *different* wave-lengths (see Figure 1), and the second is the *density* at which each of these wave-lengths is transmitted. The saturation was, therefore, calculated in two ways; first, the total number of different wave-lengths transmitted by each filter regardless of the density was computed, and, secondly, the product of the wave-lengths and their density was calculated. In calculating the saturation by the second method, only those wave-lengths which deviated from the mode by *more* than 20 mμ were included, i.e., all wave-lengths falling within 20 mμ of the modal wave-length were disregarded, since no filter is absolutely monochromatic. It was found that it makes no difference which method of measuring the saturation is used, for a perfect correlation between the two exists.

The correlation between the saturation difference and the difference between the members of the pairs as measured in j.n.d.'s is .782 and as measured by wave-lengths is .66. This is what might be expected since the correlation between the saturation and the wave-length of the individual filters is .90. It seems, therefore, unlikely that the correlations obtained between the measures of saturation differences and either mixture or dominance differences would represent *only* the effect which saturation has upon the binocular field. No correlations between saturation and either dominance differences or mixture are offered in substantiation of the suggestion that the saturation of the observed colors influences the binocular color. Were there enough cases to permit the calculation of the correlation by the product-moment method the effect which is due to wave-length difference might be partialled out and the effect of the chroma difference left. With only 10 cases such a method would be too unreliable to justify any conclusions based upon it.

In view of these facts, the effects of saturation and wave-length seem to be too inextricably mixed to permit an exact determination of either. With the facts which have just been deduced as to saturation, it seems likely that the dominance differences are in part, at least, due to differences in saturation as well as to wave-length differences.

B. Effect of Certain Physiological Conditions upon the Phenomena Reported

The preceding sections have shown that various aspects of the stimulus will be influential in determining the color of the binocular field. It will be the problem of the succeeding sections to show that certain physiological conditions also have a marked effect upon its color and brightness.

1. *Effect of Adaptation*

a. Change in brightness. Every presentation in this experiment was made to eyes in a state of light adaptation (see "Procedure") and was, with certain specified exceptions, of 58 seconds' duration. The observation was made with a light-adapted rather than with a dark-adapted eye because of the rapidity with which the eye adapts to light during the first few seconds of exposure; Geldard (15) found that after the first rapid loss of sensitivity a place is reached where the decrease in sensitivity is much more gradual. In this experiment, however, another factor was introduced. When the eye is stimulated by selected groups of wave-lengths (hues), instead of by the total of all wave-lengths (white light), it is found that the adaptation process is more rapid for certain wave-lengths than for others. The difficulty of interpretation here lies in the fact that, in much of the earlier work, the problem was neither adequately controlled nor approached from a quantitative angle, with the result that the reported data are very conflicting.

Schön (36) reported an experiment which was an attempt at quantitative measurement, though the crude apparatus made the results of doubtful reliability. He found that adaptation per unit of time was least for *R* and greatest for *B*. Wanach (45), on the contrary, found that the less refrangible rays adapt more quickly and suggested that the rapidity of adaptation is directly proportional to the length of the wave. The most recent and carefully controlled work on this subject is that done by Almack (1). The results, however, are reported in energy units, making them difficult of interpretation for our problem; the time of stimulation is plotted against the log of the reciprocal of the energy (in watts x 10^{-13}) necessary to arouse a threshold sensation. In these terms the *R* adapted most rapidly, with *B, G,* and *Y* following in that order; the colors used were photometrically equal. There is no information given which enables us to evaluate and compare the energy units with the photometric units. The technique of both Schön and Wanach was to compare a color aroused upon an adapted part of the retina with the same color upon an unadapted part of the same retina; a photometric match was then made. Had Almack given the color temperature of her lights, it would be a simple matter to find the corresponding brightnesses after adaptation and consequently the relationship of her results to those of the present study.

When *B* and *R*, e.g., were observed together, during the first few seconds a very rapid alternation of dominance occurred, with the *R* and *B* about equal in dominating power; the fluctuations soon became slower, however, and the periods over which *R* dominated tended to become longer and longer, so that the dominance difference was almost invariably positive for this pair. Since the pairs in which the two members are less widely separated in wave-length show no such marked change in dominance during the course of a single observation, it is suggested that the rapidity of adaptation is a direct function of the wave-length, although the conditions under which these observations were made are so complex as to make this suggestion little more than speculation.

This apparent differential color-adaptation introduces another problem. If different hues show different adaptation rates it appears logical to suppose that the length of time required for the eye to regain a specified equilibrium is dependent on the level to which the previous stimulation has reduced it. Data regarding the rate at which an eye recovers from adaptation to colored lights are absent from the literature, so the minimum time allowed between each observation was arbitrarily set at one minute. Since a few of the pairs tended to show a regular decrease in the dominance of one color and an increase in the dominance of the other during the five observations, it would seem that in these cases the time was inadequate. It was felt safer, however, to keep the time a constant than to introduce an unknown variable by changing it in these cases.

These data indicate that with adaptation the longer wave-length becomes brighter (comparatively) than the shorter and consequently tends to dominate more, which corroborates the evidence deduced from the observations of the colors at different physical intensities.

b. Changes in hue. Besides the change which adaptation produces in brightness, hue changes also. Hess (17) reported that all colors had a *B* or *Y* valency and that, upon adaptation, every color tended to shift toward the one which lay nearer to it in wave-length. Voeste (44) found essentially the same thing, but reported his data in rather more quantitative form.

An inspection of individual records shows that, in cases where the colors mix only part of the time, rivalry occurs at the beginning of the observation, while mixture develops toward the end; the reverse is never the case. From this we may conclude that here mixture is largely dependent on the degree of brightness adaptation and the direction in which the hues of the two colors shift. If the pairs of colors are examined in light of the shifts in hue which Hess and Voeste report as occurring with adaptation, we find that in those cases where mixture occurs, the shift in the hues is toward colors which are non-complementary, i.e., both colors shift toward *Y* or toward *B*, while in those which fail to mix the shift is toward complementary colors, i.e., one color shifts toward *B* and one toward *Y*. We are implying here that those colors which mix are aided in doing so by a shift in their *apparent* wave-lengths, which become, with adaptation, more nearly alike, and consequently produce the same phenomenon as is produced by two color filters which transmit to some extent identical wave-lengths. Those colors which do not mix, on the contrary, shift away from each other with adaptation, producing the effect of transmitting entirely heterogeneous, and therefore non-mixing wave-lengths.

In the 10 pairs of colors there are three apparent exceptions to this rule, *B* and *YG*, and *BG* and *YG* which do mix, and *R* and *YG* which sometimes do and sometimes do not mix. The case of the last pair can be most readily disposed of. This pair did not mix for C and L, but was consistently reported by *G* as mixing to produce *Y*. This pair might have been expected to produce *Y* for two reasons: (*a*) with adaptation both colors shift toward *Y*, and (*b*) the colors are nearly, though not quite, complementaries, and on the basis of the Young-Helmholtz theory might be expected to produce *Y*, and on the basis of the Hering theory would cancel to leave residual *Y*. Both C and L reported only rivalry, but they agreed that, as the observation continued, the *G* became decidedly yellowish; reports made specifically on this point show that neither C nor L saw the color as sufficiently smooth and stable to report mixture. Possibly had the observation continued longer these *O*'s would also have reported mixture, as, indeed, C did for brief flashes after 2 minutes of continued observation. Since the distance separating *R* and *YG* is great (170 mμ or 64 j.n.d.'s) it is quite possible that 58 seconds' adaptation time could not bring the two colors near enough together in hue to produce apparent overlapping of the wave-lengths and consequent mixture. In the case of the other two pairs, we find that both mix, although the shift of the one component hue, with adaptation, is toward *B* and of the other, toward *Y*. The distance between the *B* and *YG*, however, is only 90 mμ or 50 j.n.d.'s. *BG* and *YG* are even nearer together, 40 mμ or 23 j.n.d.'s separating them. This rather small separation, especially in the latter case, where the two filters transmit a large number of identical wave-lengths, may produce too much homogeneity between the filters to be overcome in the 58 seconds' adaptation time.

2. *Effect of Differential Accommodation.* As early as 1874 Bezold (3), working on the problem of binocular mixture, attributed his failure to get certain colors to mix to the fact that the eye accommodates differently to wave-lengths of different degrees of refrangibility. In his apparatus the colors were set at the same distance from the eye so that, because of such differences in accommodation, one of the uniocular images was necessarily out of focus. He therefore suggested that the phenomenon of rivalry was closely correlated with fluctuations of accommodation, the dominant field being the one in focus. Dobrowolsky (8) repeated Bezold's work the next year (1875), using an apparatus to correct for this accommodation difference, and in general confirmed his conclusions. The slight fluctuations that did occur he attributed to fluctuations of convergence, since he had controlled accommodation.

In order to study the effect of accommodation on our results, the apparatus was arranged so as to make possible a differential adjustment of the fixation points. The optimal location of each fixation point was approximately determined by the method of limits for two *O*'s (C and G). With the fixation points in focus a complete series of observations was made on both C and G. Table 5 gives the figures for dominance difference and mixture for both *a* and *β* series for each of the two *O*'s; no very significant changes result. For G the positive dominance differences are slightly but consistently increased; for C the same is true, though it is less marked. The difference in the effect on the two *O*'s may be explained partly by the fact that for C the fixation points as they were placed in the original series more nearly approximated the optimum than for G.

A comparison of the correlations (Table 3) of this series and the

TABLE 5

Effect of Various Physiological Factors, as Shown by the Size of the Dominance Difference and the Amount of Mixture Reported

	Differential accommodation				Blinking				No blinking		Three controls	
	Obs. C		Obs. G		Obs. G		Obs. L		Obs. L		Obs. C	
Pair	D.D	Mix.	D.D.	Mix.	D.D.	Mix.	D.D.	Mix.	D.D.	Mix.	D.D.	Mix.
Alpha Series												
1	....	49.0	....	49.0	....	49.0	....	49.0	....	49.0	....	49.0
2	23.3	19.9	39.0	6.0	41.6	....	....	47.2	13.8	34.4	16.6	32.4
3	....	48.0	....	48.8	....	49.0	....	49.0	....	49.0	....	49.0
4	44.5	....	43.7	....	41.2	....	–.7	11.2	18.3	....	41.2	5.7
5	–1.9	33.8	....	49.0	–.2	45.2	37.2	....	31.5	....	–5.2	22.7
6	....	48.7	44.2	....	35.0	....	21.5	....	23.4	17.0	26.4	14.9
7	24.3	....	.6	44.0	–17.2	28.2	29.6	....	.5	....	22.7	....
8	46.4	....	43.9	....	28.1	....	16.6	....	20.6	....	45.2	....
9	34.5	....	17.9	....	11.9	....	23.5	....	3.3	....	40.9	....
10	27.8	....	20.0	....	19.2	....	20.3	....	23.5	....	27.9	....
Beta Series												
1	....	47.8	....	48.8	....	49.0	....	49.0	....	49.0	....	49.0
2	12.8	36.0	....	49.0	33.5	4.9	....	47.0	7.8	40.4	19.7	29.0
3	....	49.0	....	49.0	....	49.0	....	49.0	....	49.0	....	49.0
4	43.9	....	49.5	....	31.7	....	–25.7	8.9	–3.6	....	36.8	....
5	–2.1	11.4	....	48.8	–6.7	42.4	30.1	....	14.2	....	–9.7	5.2
6	22.6	26.2	46.2	....	28.3	10.0	–7.6	....	–8.8	4.6	29.2	.3
7	15.6	....	–16.4	30.7	–13.0	36.1	.8	....	–2.6	....	25.5	....
8	49.0	....	35.7	....	34.5	....	–20.2	....	–19.4	....	42.8	....
9	30.5	....	28.1	....	7.8	....	15.1	....	3.9	....	33.7	....
10	34.2	....	27.6	....	18.7	....	–12.7	....	–8.4	....	15.7	....

original series indicates that the differences between the two are not great. In general, the correlations for dominance difference are slightly increased in a positive direction, while those for mixture are consistently decreased.

The difference which optimal accommodation conditions made in the reports of our *O*'s does not seem to be as great as the work of the earlier investigators would lead us to expect. With a single exception, neither *O* reported mixture where none had previously been reported; both remarked on the fact that the observation was perhaps slightly less of a strain. From C's introspection regarding the one exceptional mixture (*B* and *R* producing *V*, after 3 minutes' observation), it appears that the standard for mixture set up for our *O*'s was much stricter than Dobrowolsky's and the apparent importance of accommodation is thereby greatly reduced.

3. *Effect of Blinking.* The amount of blinking which occurred

during an observation was considered as a possible source of variability in the records. This was controlled in a complete series for each of two *O*'s (L and G) by having *O* blink in time to a metronome, whose bell rang every 4 seconds. For L a counter-series was run in which the instructions were not to blink at all during the observation; he reported only an occasional slip in following these instructions. A comparison of the absolute figures for the "metronome" series with those for the "no-blinking" series shows no significant difference. (Table 5.) A consideration of the correlations in these series leads one to assume, however, that for L the fairly rapid and accurately timed blinking tends to make and keep the longer wave-length dominant, while for G rapid blinking tends to bring back the weaker color, which was usually the shorter wave-length. This is borne out by voluntary introspective reports on the part of both *O*'s. (Table 3.)

C did not observe in a series given for the sole purpose of testing the effect of blinking, but a series was taken with him in which the fixation points were set at their optimal focal distances, the period allowed for recovery slightly lengthened, and the metronome used to control blinking. In this series we find that, compared to the series with the fixation points in focus, there are possibly very slightly smaller dominance differences. (Compare results for C, Table 5, "Differential Accommodation" and "Three Controls.") The correlations also indicate that for C, like G, the regular blinking tends to bring back the color which had been suppressed, i.e., usually the shorter wave-length. (Table 3.)

4. *Effect of Previous Uniocular Fatigue.* The apparently important rôle which adaptation to colors plays in this problem made it advisable to determine the effect of adapting the eye to one of the colors before making a record in the usual way.

For the purpose of studying this effect the electrical connections of the lights were readjusted. By the new arrangement it was possible to turn either light on first, having it burn at the appropriate voltage, and then switch in the second light after any desired interval. The method used was to adapt the eye to one member of a pair for one minute and then turn on the second light; the two then continued to burn for 58 seconds while the record was made in the usual manner. C observed through a complete *a* series of this type. (Table 6.) Between each observation sufficient time was

TABLE 6

THE EFFECT OF PREVIOUS UNIOCULAR ADAPTATION TO ONE MEMBER OF THE PAIR

Observer C — ALPHA SERIES

Pair	60″ adaptation to long wave-length D.D.	60″ adaptation to long wave-length Mix.	60″ adaptation to shorter wave-length D.D.	60″ adaptation to shorter wave-length Mix.
1	.....	30.5	10.75	38.25
2	−5.75	27.25	7.75	40.75
3	−4.0	43.5	31.75	17.0
4	34.25	.....	23.5	25.25
5	0	13.25	20.75	22.25
6	−3.25	26.25	26.0	23.0
7	16.5	.....	44.0	.....
8	7.5	.....	48.5	.....
9	26.0	.....	41.0	.....
10	23.5	.....	37.5	.....

allowed for all after-images to completely disappear, but with a minimum of 5 minutes. Two records were made of each pair of colors with the right eye and two with the left eye previously adapted, the results being the average of these two readings. The fixation points were set at the previously determined optimal focal distances.

The results show that when the eye is adapted to the longer wave-length there is a tendency for the shorter wave-length to dominate at the beginning of the binocular observation, though it quickly drops back to its normal position of having less dominating quality. In the case of the pairs that usually mix from the beginning of an observation there is a period of fluctuation in which either the shorter wave-length dominates, or in which neither can be reported as dominant. When the eye is previously adapted to the shorter wave-length the tendency is to emphasize the positive dominance differences. The pairs that usually mix completely from the beginning now show a dominance of the longer wave-length before the mixture develops, and in all cases the shorter wave-length practically fails to appear. In every case the color seen by the adapted eye grew dull and was reported as either completely disappearing or at least fading out very definitely when the second color was turned on. The other color appeared bright and saturated. As the observation of both colors continued, an equilibrium was attained and the last fourth of the record indicates, what the introspective reports confirm, that the two colors appear, relatively to each other, as they do in the regular observation.

These results show the tendency of the shorter wave-lengths to adapt more quickly and completely than the longer ones; this tendency, as earlier suggested, may account in large measure for the greater dominance of the longer wave-lengths.

Summary of Results

The results of stimulating the two eyes simultaneously with different colors by the method described may be summarized as follows:

1. The color of the member of the pair which has the longer wave-length dominates the binocular field more than that of the shorter.
2. Those pairs of colors, the members of which are separated by the smallest number of wave-lengths or j.n.d.'s, mix the most readily and completely.
3. The nearer together the two colors in wave-length, the fewer the fluctuations.
4. The color having the greater intensity has the greater dominating power.
5. The less the absolute brightness of the colors, the more readily they mix.
6. The more alike in brightness the colors, the more readily they mix.
7. With adaptation, the shorter wave-lengths tend to lose more in brightness than the longer.
8. When the fixation points are in focus, the positive dominance difference tends to be increased.
9. Regular and fairly rapid blinking tends to bring back the weaker color, though individual differences here are large.
10. When the eye is previously adapted to the color having the shorter wave-length of a pair, the dominance difference is increased in a positive direction; adaptation to the longer wave-length tends to decrease the size of the dominance difference.

References

1. Almack, M. R. A quantitative study of chromatic adaptation. *Psychol. Monog.,* 1928, **38**. No. 2. Pp. 118.
2. Aubert, H. Physiologie der Netzhaut. Breslau: Morganstern, 1865. Pp. 394.

3. Bezold, W. v. Ueber binoculare Farbenmischung. *Pogg. Ann. d. Physik u. Chemie* (Jubelband), 1874, 585.
4. Breese, B. B. Binocular rivalry. *Psychol. Rev.*, 1909, **16**, 410-415.
5. ———. On inhibition. *Psychol. Rev., Monog. Suppl.*, 1899-1901, **3**. Pp. 65 (esp. 20 ff.)
6. Chaúveau, A. Sur la théorie de l'antagonisme des champs visuels. *Comp. rend.*, 1891, **113**, 358, 361, 439 ff.
7. Dawson, S. Experimental study of binocular color mixture. *Brit. J. Psychol.*, 1915-1917, **8**, 510-551.
8. Dobrowolsky, W. Ueber binoculare Farbenmischung. *Pflüg. Arch. f. d. ges. Physiol.*, 1875, **10**, 56-62.
9. Dove, H. W. Ueber Binocularsehen durch verschiedene gefärbten Gläser. *Pogg. Ann. d. Physik u. Chemie,* 1857, **101**, 147-151.
10. ———. Ueber das Binocularsehen prismatisher Farben und eine neue stereoskoptische Methode. *Pogg. Ann. d. Physik u. Chemie,* 1850, **80**, 446-448.
11. ———. Ueber Binocularsehen und subjective Farben. *Pogg. Ann. d. Physik u. Chemie,* 1861, **114**, 163-168.
12. Downey, J. E. Types of dextrality and their implications. *Amer. J. Psychol.,* 1927, **38**, 317-367.
13. Fechner, G. T. Ueber einige Verhältnisse des binocularen Sehens. *Abh. d. Sächs. Ges. d. Wissen.*, 1861, **7**, 391, 397, 384, 388.
14. Frey, N. v., & Kries, J. v. Ueber die Mischung von Spektralfarben. *Arch. f. Anat. u. Physiol.* (Physiol. Abteil.), 1881, 336-353.
15. Geldard, F. A. The measurement of retinal fatigue to achromatic stimulation. I & II. *J. Gen. Psychol.*, 1928, **1**, 123-135; 578-590.
16. Haldat, C. N. A. Sur la double vision. *J. de physique,* 1806, **63**, 387.
17. Helmholtz, H. v. Physiological optics. (Eng. trans.) Optical Society of America, 1925.
18. Henning, H. Ein optisches Hintereinander und Ineinander. *Zsch. f. Psychol.*, 1921, **86**, 144-174.
19. Hering, E. Beiträge zur Physiologie. Leipzig: Engelmann, 1861. Vol. 5, 308.
20. Hess, —. Ueber die Tonänderungen der Spektralfarben durch Ermüdung der Netzhaut mit homogenem Lichte. *Graefes Arch. f. Ophthal.*, 1890, **36**, 1-33.
21. Jones, L. A. The fundamental scale of pure hue and retinal sensibility to hue differences. *J. Opt. Soc. Amer.*, 1917, **1**, 63-77.
22. König, A. Ueber die Anzahl der unterscheidbaren Spektralfarben und Helligkeitsstufen. *Zsch. f. Psychol.*, 1895, **8**, 375-400.
23. ———. Ueber den Helligkeitswert der Spektralfarben bei verscheidener absoluter Intensität. *Zsch. f. Psychol.*, 1891, (Helmholtz Festschrift), 309-388.
24. König, A., & Dieterici, C. Ueber die Empfindlichkeit des normalen Auges für Wellenlängenunterschiede des Lichtes. *Pogg. Ann. d. Physik u. Chemie,* 1884, (n. F.), **22**, 579-589.
25. Kuroda, G. An experimental study of binocular mixture. *Jap. J. Psychol.,* 1923, **3**.

26. ————. Investigations on the temporal course of rivalry of the visual fields with different colors. (Japanese.) *Kioto-Zgaku-Zasshi,* 1915, **12**.

27. ————. Objective conditions of binocular predominance, rivalry, and mixture. *Jap. J. Psychol.,* 1919, **1**.

28. ————. Zur Grenzbestimmung der binokulare Phänomene. *Psychol. Forsch.,* 1925, **6**, 282-296.

29. LANGE, N. Beiträge zur Theorie der sinnlichen Aufmerksamkeit und der activen Apperception. *Phil. Stud.,* 1888, **4**, 390-422. (Esp. 417 ff.)

30. LAURENS, H., & HAMILTON, W. F. The sensibility of the eye to differences in wave-length. *Amer. J. Physiol.,* 1923, **65**, 547-568.

31. NUTTING, P. G. The reactions of the retina to light. *Amer. J. Physiol. Opt.,* 1920, **1**, 142-170.

32. PANUM, P. L. Physiologische Untersuchungen über das Sehen mit zwei Augen. Kiel: Schwers, 1858. Pp. 94.

33. RIVERS, W. H. R. On binocular color mixture. *Proc. Cambridge Phil. Soc.,* 1895, **8**, 273-277.

34. ROCHAT, G. F. Étude quantitative du fusionnement binoculaire des couleurs complementaires. *Arch. néerl. de physiol. d'homme & des anim.,* 1922, **7**, 263-267.

35. ROELOFFS, C. O. Ueber Wettstreit und Schwankungen im Sehfelde. *Arch. f. Ophthal.,* 1921, **104**, 230-263.

36. SCHÖN, W. Einfluss der Ermüdung auf die Farbenempfindungen *Arch. f. Ophthal.,* 1874, **20**, 273.

37. ————. Zum Wettstreit der Sehfelder. *Klin. Monatbl. f. Augenheilk.,* —, **8**, 356.

38. ————. Zur Lehre vom binocularen Sehen. *Arch. f. Ophthal.,* —, **24**, 66.

39. SMITH, F. O. An experimental study on retinal sensitivity and discrimination for purple under different degrees of intensity of stimulation. *J. Exper. Psychol.,* 1925, **8**, 381-397.

40. STEINDLER, O. Die Farbenempfindlichkeit des normalen und farbenblinden Auges. *Sitzungsber d. Akad. d. Wissensch. z. Wien: Math.-Naturw. Klasse,* 1906, 115, (2. Abt. A), 39-62.

41. TRENDELENBURG, W. Versuche über binoculare Mischung von Spektralfarben. *Zsch. f. Psychol.,* 1913, **48**, 199-210.

42. ————. Weitere Versuche über binoculare Mischung von Spektralfarben. *Pflüg. Arch f. d. ges. Physiol.,* 1923, **201**, 235-246.

43. TROLAND, L. T. Colors produced by equilibrium photopic adaptation. *J. Exper. Psychol.,* 1921, **4**, 344-390.

44. VOESTE, H. Messende Versuche über die Qualitätsänderungen der Spektralfarben in Folge von Ermüdung der Netzhaut. *Zsch. f. Psychol.,* 1898, **18**, 258-267.

45. WANACH, B. Eine Notiz über Farbenermüdung. *Zsch. f. Psychol.,* 1909, **43**, 443-446.

46. WOINOW, —. Ueber den Wettstreit der Sehfelder. *Arch. f. Ophthal.,* —, **16**, 194.

University of Rochester
Rochester, New York

UNE ÉTUDE QUANTITATIVE DE LA VISION BINOCULAIRE DE COULEURS

(Résumé)

Le problème de cette expérience a été l'investigation des conditions physiques qui déterminent les phénomènes présents quand l'on observe en même temps deux couleurs, une avec chaque oeil. En employant des filtres de gélatine approximativement monochromatiques comme stimuli, on a présente en même temps des paires de couleurs au moyen d'un appareil contrôlant l'intensité et empêchant l'entrée de lumière extérieure. On a trouvé qu'avec des intensités égales de lumière sur les deux yeux, les plus grandes longueurs d'onde dominent la plupart du temps au champ binoculaire; quand les intensités sont inégales, la plus forte tend à dominer plus. Un mélange des deux couleurs composantes en un mélange parfaitement homogène et stable, est arrivé seulement quand les deux filtres de la paire employée se sont superposés un peu dans la rangée des longueurs d'onde transmises; les couleurs d'intensités égales mais basses se sont mélangées le plus facilement. Le nombre de fluctuations pour chaque unité de temps a été une fonction de la distance entre les couleurs composantes. On a trouvé aussi qu'outre divers aspects du stimulus qui influent sur le champ binoculaire, de certaines conditions physiologiques sont des facteurs déterminants. Quand les points de fixation ont été au foyer, la dominance de la plus grande longueur d'onde est devenue plus grande. Avec adaptation les plus courtes longueurs d'onde ont paru perdre plus de clarté que les plus grandes, ce qui a semblé un facteur dans la détermination de la qualité dominante supérieure de la plus grande longueur d'onde. L'adaptation antérieure d'un oeil à un membre d'une paire a fait décroître la dominance de cette couleur dans l'observation subséquente de la paire de couleurs.

JOHANNSEN

EINE QUANTITATIVE UNTERSUCHUNG DER BINOKULAREN FARBENWAHRNEHMUNG

(Referat)

Die Aufgabe des gegenwärtigen Experiments war die physischen Bedingungen zu untersuchen, welche die Erscheinungen bestimmen, die beim gleichzeitigen Beobachten von zwei, dh. je einem Farbenton mit jedem Auge auftreten. Indem man nahezu monochromatische Gelatinfilter als Reize gebrauchte, gab man gleichzeitig Farbenpaare mittels eines Apparates, der die Regulation der Helligkeit gestattete und Streulicht verhinderte. Man fand, dass mit gleicher Lichthelligkeit auf beide Augen die längern Wellenlängen die meiste Zeit das binokulare Feld beherrschen; wenn die Helligkeiten aber ungleich sind, so dominiert die stärkere um so mehr. Mischungen der zwei Farbenbestandteile in eine vollkommen homogene und konstante Farbe zeigten sich nur dann, wenn die Wellenlängenbereiche der beiden Farben etwas übereinandergreifen; Farben gleicher und niedriger Helligkeit mischten sich am leichtesten. Die Zahl der Schwankungen pro Zeiteinheit erwies sich als Funktion der Distanz zwischen den Farbenkomponenten. Man fand ferner, dass neben den verschiedenen Merkmalen (aspects) der Reize, die das binokulare Feld beeinflussen, gewisse physiologische Bedingungen mitbestimmende Faktoren sind. Wenn der Fixierungspunkt auf den Brennpunkt eingestellt ist, wird das Vorherrschen der längern Wellenlänge erhöht. Mit Anpassung schien die kürzere Wellenlänge an Helligkeit mehr zu verlieren als die längere, was wohl das Vorherrschen der längern Wellenlänge bestimmt. Vorhergehende Anpassung des einen Auges auf eine der beiden Farben führte zu vermindertem Vorherrschen jener Farbe in den nachfolgenden Beobachtungen des Farbenpaares.

JOHANNSEN

CHANGES IN TONUS DURING COMPLETED AND INTERRUPTED MENTAL WORK*

From the Psychological Laboratories of the Institute of Human Relations, Yale University

G. L. FREEMAN

The experiments herein reported measured the changes in muscular tension which occurred during the solution and interruption of various tasks. Our first task was to explore, in a number of subjects, the tonus accompaniments of "mental work." With the observation of certain regularities in the records, we formulated our second experimental problem; the comparison of tonus changes during related periods of interrupted and continued performance. This experiment repeated the essential conditions of a recent research of the Berlin Laboratory (17), which is exploited by Lewin in support of his doctrine of *"psychic tensions."* Our results, while confirming those of the earlier experiment, suggest an alternative explanation. This we shall discuss in an appropriate place.

Tonus has long been considered an excellent indicator of central nervous activity. Variations in muscle tone produced by voluntary action and "attention" have been extensively studied through the medium of the knee-jerk. The recent work of Golla (10), Travis (18), Bills (4), Tuttle (19), and Jacobson (14) has demonstrated that mental activity and muscular tension are related in some intimate way. Very little is actually known concerning the conditions of this relationship, since the tonus accompaniments of mental work are hard to isolate under experimental conditions. This difficulty can be lessened by the use of a more sensitive and accurate recording technique, such as is now available in our method of tendon deformation (9).

---

*Accepted for publication by Raymond Dodge of the Editorial Board and received in the Editorial Office, March 10, 1930.

[1] This research was carried out under a National Research Fellowship in the Biological Sciences. The writer is indebted to the Institute of Human Relations for the use of excellent laboratory facilities and to Professor Dodge for many helpful suggestions and a critical interest in the experiments.

Experimental Methods and Procedures

Our choice of subjects, tasks, and method of recording muscular tension was determined by a number of factors. Certain of these were mentioned in connection with our previous discussion of tonus measurement. Other factors remain to indicate added difficulties which will be met in the exploration of the tonus–mental-work relationship. Our methods of meeting the difficulties relevant to our problems are summarized below.

A number of internal variants enter into mental work. Each of these variants may influence the level of neural-muscular tension in its own way. Since it is impossible rigidly to control such little understood conditions, as the degree of effort, motivation, difficulty of task, and general psycho-physiological condition, the use of subjects trained for commentative report is highly desirable. On the other hand, the sophisticated psychological observer presents real difficulties in certain types of experimental variation. The suggestibility of a naïve subject is greater in the same situation and his motivation may be more easily controlled. We aim to equilibrate such consequences by including the protocols of both naïve and trained subjects.

Mental work embraces two general types: That which comes to a definite end, such as solving a puzzle or a problem in calculus, and than which can go on indefinitely, such as writing. Both types may involve the continued use of a single psychological operation, the discontinuous use of a single operation, or the successive use of several different operations. The activity of the organism differs, conceivably, under these several procedures. Our experimental tasks included, therefore, samples of all these variations.

Our method of measuring tonus by patellar-tendon deformation demands an isometric quadriceps muscle. This limited the range of tasks to such as could be executed in the sitting position. Additional experimental restriction was necessary in order to secure reliability of the records. The leg had to be securely stabilized on both sides of the knee and a check obtained upon slight back-front movements of the limb. Such movements, occasioned by restless shifts of the trunk and pelvis, could, of course, be minimized by stressing a rigid posture. But it was feared that the instruction to "sit still" might create an increased general tension quite detrimental to the measurement of the normal tonus accompaniments of mental work. Another difficulty in the interpretation of tonus changes regards the question of whether the records are correlated with the spread of mental or gross muscular excitation. We have met these difficulties by checking purely mental work against tasks which involved the hands and by eliminating from our data records and readings which were accompanied by thigh displacement.

The recording device was, in essential, a lever serving to depress the patellar tendon. This lever was connected with an optical lever which magnified changes in tendon deformation approximately 500 times. A thread attached to the depressor-lever at its point of contact with the tendon encircled a small cylinder (radius 2 mm.) carrying a convex mirror of suitable curvature (the optical lever). The thread was held under constant tension by a rubber band. The center of a 180-degree screen made of cardboard was placed 50 cm. away from the axis of the cylinder. The surface of this curved screen was covered with mm. cross-section paper. After the depressor lever had been placed against the subject's tendon at the point of greatest deformability, the beam of light reflected by the optical lever was brought by a suitable adjustment to the center of the millimeter-scale. During the experimental hour no further adjustments were made, except in rare cases where leg displacement caused the beam of light to leave the scale. The experimenter simply read the level of muscular tension from the scale at stated intervals. Changes in the records produced by movement of the limb were checked by the position of the apparatus carrying the tension lever. This position was projected upon the scale by the use of a small mirror situated on the carrying apparatus. Any movement of this reflection during mental work indicated that the tendon-reading was an unreliable sign of the spread of neural excitation at that movement.

Experiment 1

This experiment compared the tonus changes occurring during mental work with preceding periods of relaxation. The level of tension in the right quadriceps muscle was obtained every 30 sec. during both working and non-working periods. The experiments were conducted in a partially darkened room. The subject sat in a morris chair. A board placed across the arms of this chair served as a table and carried a shaded electric lamp. A screen (above the board) shut off the view of the recording apparatus.

An attempt was made to accustom *S* to the experimental environment before commencing to record. After the apparatus was properly placed *E* conversed with *S* upon topics of general interest. During the initial period it was casually mentioned that some simple mental tasks would be given "after a while." By thus satisfying the curiosity of the subject, initial anticipatory tension was noticeably reduced. When later *S* was asked to "close the eyes and rest for a time," "normal" relaxation was easily recorded.

Various mental tasks were used in the several series. Maximal effort was sought in all cases by the use of appropriate incentives. These individual variations are given below.

Series 1a: Continuous Operations

Two naïve (W and B) and two trained (P and F) subjects worked upon tasks calling for the use of a single operation. Four such tasks were continued for a period of 10 minutes each, with equivalent interpolated rest periods. Two of the tasks (continuous addition, counting verbs and adjectives) did not directly necessitate muscular activity; the other two (writing *a*'s, cancelling *f*'s) did. The *S*'s performed all four tasks during the same hour but in different time orders. Each task thus occurred at the beginning, middle, and end of the experimental period. Rivalry was suggested as an incentive to maximal effort. A sample of the formal instruction, such as preceded each of the four tasks is given below.

> "You are not to be tested upon the speed with which you can count adjectives and verbs. When I say ready open your eyes and begin reading the selection before you. Work until I tell you to stop. After you have told me the number of adjectives and verbs you were able to count, indicate the place where you ceased reading. Then close your eyes and rest quietly."

Information was sought by free and also by direct report upon the following questions:

1. "Was the reading of the formal instruction followed immediately by anticipatory thought of the task?"
2. "Was the same degree of effort expended at the beginning as at the end?"
3. "What was the motivation?"
4. "Did the subject tend to 'think about' a previous task or the next one during the interpolated rest-period?"

Series 1b: Discontinuous Operations

Three trained (P, L, and Se) and three naïve (T, Su, and O) subjects served in the experiment. Each completed a series of 30 trials in one hour. *S* reacted to the seen movement of a small visual object by a manual movement in the opposite direction. A small shutter of the screen in front of *S* opened to expose two bright dots (pin heads) on a darkened background. The right-hand dot was movable. The left dot was fixed. When the right dot moved up the subject pulled down on a lever and vice versa. The stimulus and reaction system were arranged for protochronographic recording, the mechanics of which need not concern us here. In this series the changes in tendon deformation were also recorded photographically. The desirability of this sensitive recorder for acts of short duration calls for no comment. The camera, carrying a roll of continuous film, was placed at the same distance from the cylinder as was the millimeter scale of Series 1*a*.

The subjects were given an initial 10-minute period of relaxation before the specific instructions were read. Approximately 30 seconds intervened between each of the 30 reactions. Maximal effort was again sought by suggesting that "those who do best will be retained for later work and will be paid well for their services." The instructions serving to initiate the reaction performances were:

> "When I say ready, open your eyes and fixate the two dots which dominate the visual field. If the right-hand dot moves upward, give the stick in your hand a quick downward pull. If the dot moves down push up with the stick. React quickly and accurately. Close your eyes after each reaction and await the ready signal for the next trial."

Series 1c: Alternating Operations

Three naïve subjects (T, Su, and O) were each given three sets of calculations. Each set involved the use of all four simple arithmetic operations and was planned to require about 10 minutes for completion. Ten minutes of relaxation preceded each set of calculations. Changes in tendon deformation were read at 30-second intervals on the millimeter scale as in Series 1*a*. Reports taken at the close of the tests sought information upon the questions similar to those advanced in the preceding series. The incentive to maximal effort was the hope of reward.

Three trained graduate students in psychology (Se, Sa, and M) were each given an equivalent number of psychology tests. The tests covered (*a*) theoretical psychology, (*b*) abnormal psychology, and (c) physiological psychology. We sought to make experimental conditions approximate those given to the three naïve subjects, as stated above. The tests were of the true-false variety and involved the alternative use of memorial, comprehensive, and elaborative operations. Rivalry furnished an incentive to maximal effort.

EXPERIMENT 2

These experiments compared the tonus-changes occurring in equivalent periods of interrupted and uninterrupted mental work. The types of mental work and interruptions approximated those used in the Lewin experiment. Two series of 10 simple problems were given to each of 10 subjects in the same experimental hour. The tasks of the first series could be characterized as purely mental. The tasks of the second series involved manual movements of writing and puzzle solving. The test averaged 5 minutes for completion. A period of approximately 2 minutes intervened between successive tasks, during which the subject sat with eyes closed.

In the case of each *S*, five of the problems in both series were interrupted one minute after they were begun. The interruptions were of two types: (*a*) accidental interruption (*A.I.*), such as the automatic cutting off of illumination, and (*b*) deliberate interruptions (*D.I.*), such as the request from the experimenter to begin something new. No matter how serious the interruption, opportunity was presented for the resumption of the old task during the minute following the initial disturbance. The schedule below shows how all tasks received both accidental (*C*) and deliberate (*D*) interruptions.

The presentation of accidental and deliberate interruptions required an extremely deft handling of the experimental situation. Sophisticated subjects are likely to obtain an insight into the experimenter's purpose almost at once. Even naïve individuals tend to "ignore disturbance" when it becomes habitual. A serious effort was made, therefore, to have the interruptions appear both natural and unintentional. The interruptions were of the following order.

Accidental Interruptions

A_1. A large piece of cloth which was suspended above the chair fell on *S*'s head. With appropriate apologies and explanation, the experimenter soon succeeded in lifting the cloth by use of a convenient stick. In its original position the cloth suggested a photographic hood.

A_2. The subject's desk lamp was tipped against him by pulling on a concealed string. The experimenter finally rearranged the lamp

TABLE 1

SCHEDULE FOR THE PRESENTATION OF TASKS AND INTERRUPTIONS IN EXPERIMENT 2

		Time order 1-2*					Time order 2-1				
	Tasks Subj.:	Se	R	Su	Fa	T	Sa	Y	L	G	W
	Series I										
1.	Continuous addition	C_1		D_1		C_2		C_3		D_2	
2.	Continuous subtraction		D_2		C_3		C_2		D_1		C_1
3.	Counting verbs and adjectives	D_1		C_2		C_3		D_2		C_1	
4.	Combined addition and subtraction		C_3		C_2		D_1		C_1		D_2
5.	Counting hidden figures	C_2		C_3		D_2		C_1		D_2	
6.	Combined multiplication and division	C_3		D_2		C_1		D_1		C_2	
7.	Counting numbers ending in 3 and 7 between limits		C_2		D_1		C_1		D_2		C_3
8.	Counting *a*'s		D_1		C_1		D_2		C_3		C_2
9.	Combination arithmetic problems		C_1		D_2		C_3		C		D_1
10.	Counting syllable combinations	D_2		C_1		D_1		C_2		C_3	
	Series II										
1.	Solving jig-saw puzzle	D_3		C_4		D_4		C_5		D_5	
2.	Classification test		D_4		C_5		D_4		C_4		D_3
3.	Opposites test	C_4		D_4		C_5		D_5		D_3	
4.	Drawing simple plan		C_5		D_4		C_4		D_3		D_5
5.	Writing synonyms		D_4		C_4		D_3		D_5		C_5
6.	Writing numbers backwards	D_4		C_5		D_5		D_3		C_4	
7.	Solving ring puzzle		C_4		D_3		D_5		C_5		D_4
8.	Writing alphabet backwards		D_3		D_5		C_5		D_4		C_4
9.	General information test	C_5		D_5		D_3		C_4		D_4	
10.	Noting absurdities	D_5		D_3		C_4		C_4		C_5	

*Time order 1-2—Series I, then Series II; time order 2-1, the reverse.

and expressed regret that it had not been placed more securely on the writing board.

*A*₃. Small screws were concealed in a box located on that part of the apparatus frame which extended above the subject. By releasing a noiseless spring, these screws were "rained" down upon the subject. Apologies were again offered and the disturbance continued by picking up a few of the screws lying about the subject.

*A*₄. The current to the subject's desk lamp was cut off by throwing a silent foot-switch. During the next moment the experimenter made valiant attempts to "fix" the defective circuit. The light flashed on and off momentarily until the connection was "supposed" to be repaired.

*A*₅. Interruption A_4 was repeated.

Deliberate Interruptions

D_1. The subject was asked if he had felt himself sitting on a screw driver. Had he seen it when he entered the room? It was necessary to have it at once. An animated monologue concerning the disappearance of the object was carried on in an attempt to continue the distraction.

D_2. The subject was asked to loan the experimenter his pencil. (This usually produced an interruption of high order. Many subjects showed no tendency to surrender the object until the experimenter reached for it.)

D_3. The experimenter wondered if he had given the right problem to the subject. In order to satisfy himself, he appropriated the subject's paper and studied it for a time before deciding that it was really the correct task.

D_4. The subject was asked to write a report contrasting the difficulty of two previous tasks.

D_5. The subject was told he was not doing very well with the task at hand and that he had better do "this other problem."

It was realized that all the above changes in situation would condition organic adjustment even when not serving to disrupt an assigned task. But a means was at hand for ascertaining whether or not our several "interruptions" were accompanied by equivalent tonus change when presented independently. One subject in each time order was given the several interruptions of his second series during the relaxation periods of his first series. While the effectiveness of the interruptions was diminished at the second presentation, their average influence on tension was greater during mental work than during relaxation.

Readings of tendon deformation were taken at 30-second intervals during the first 3 minutes of task solution. Although we were interested primarily in the changes occurring during the second minute (period of possible interruption), the inclusion of readings on both sides of this period aided interpretation. The reflection which indicated the presence of gross body movement was carefully watched during the period of interruption. Fortunately, very few instances of such movement were noted. This is probably due in part to the fact that no interruptions were of a violent character and in part to a remembrance of the instruction to remain as quiet as possible during tests.

TABLE 2
Tonus Changes Occurring in Equivalent Periods of Relaxation and Continuous Mental Work

		Task 1		Task 2		Task 3		Task 4	
S	T.O.	Av. R-Change	Av. W-Change	Av. R-Change	Av. W-Change	Av. R-Change	Av. W-Change	Av. R-Change	Av. W-Change
P	1,2,3,4	+1.2	+15.40,—3.8	+ .5	+11.10,—2.0	— .8	+ 9.30,+3.1	— .2	+ 9.5,—2.4
F	4,3,2,1	+ .9	+13.20,—2.9	—1.0	+10.60,— .7	— .1	+18.20,—4.2	+ .8	+24.5,—2.9
W	2,1,4,3	+ .2	+35.70,—2.6	+3.0	+33.40,—4.2	— .9	+12.00,—2.7	— .6	+29.5,—1.8
B	3,4,1,2	+1.5	+ 8.60,—4.1	— .8	+15.20,— .3	—5.3	+ 8.20,—3.6	—3.0	+10.5,—2.9
Av.		+ .95	+18.22,—3.35	+ .93	+17.58,—1.8	—1.78	+11.93,—1.85	— .75	+18.5,—2.5

Rivalry was again suggested as an incentive to maximal effort. The formal instructions were:

> "You are to be tested upon your ability to solve various problems. Work rapidly and accurately that your score may excel those of your competitors. The tests are divided into two parts. In Part I, open your eyes with every ready signal and solve the problem which is placed before you. Do not use your hands as an aid to counting, etc. Give aloud your answer the moment you finish. In Part II, the tasks will involve the use of your hands. Close your eyes between tests as in Part I. Begin with the ready signal and indicate when you are through by tapping your pencil. Avoid shifting around during the periods of task solution. Shifts from uncomfortable positions should be made between tests if necessary."

THE RESULTS

The initiation of mental tasks under an incentive to maximal effort is almost invariably accompanied by an increase in muscular tension. Experiments conducted on a variety of subjects of task situations support the preliminary observations of Golla (10) on the positive correlation of tonus with "mental" effort. This increased spread of neural excitation due to mental work has been exhibited in a number of related studies; notably, the work of Dodge (6) on the pulse rate, of Ilzholfer (12) on the depth and rate of breathing, and that of Gopalaswami (11) on the galvanic phenomenon. We shall also present unequivocal evidence that this spread of neural excitation declines during the period of task solution. Many suggestions of this result will also be found in previous work. When a task is interrupted, tension is noticeably increased. With resumption of the task, there usually occurs a decrease in tonus. The theoretical bearing of these observations will be considered later in this paper.

Series 1a: Continuous Operations

The preceding table summarizes the tonus changes occurring in the solution of continuous tasks. It was found from inspection of the data that an initial rise in tonus regularly occurred during the first minute after the ready signal was given. We have, therefore, taken this period for estimating the average changes incident to the undertaking of the task. The three measures tabulated under each task in Table 2 are, respectively, the average millimeters of tonus change per 30 seconds for (*a*) the period of relaxation (10 minutes), (*b*) the period of task initiation (1 minute), and (*c*) the remaining period of task solution (9 minutes).

The figures given in Table 2 are the average of the individual 30-second increases and decreases of tendon deformation (as expressed in mm. deflection at the scale). These individual changes were reckoned as follows: Each reading was algebraically subtracted from the preceding. If the direction of change was positive, it was called plus; if negative, minus. That a certain mathematical relation holds between these magnified changes in tendon deformation and the amount of muscular tension was pointed out in our preliminary exploration of this method of measuring tonus. But the factors which we sought to explore in this experiment may be presented in terms of deflection changes more readily than in terms of grams tension. The successful interpolation of tension for deflection values involves an extremely difficult calibration for each individual subject. This calibration is technically impracticable for any other than photographic recording systems.

While our quantitative presentation masks certain qualitative variations, it is fairly representative of the facts. The average initial increases in tonus are lessened by the fact that the last reading for the first minute usually occurs when the release from tension has already begun. Furthermore, this presentation implies a greater regularity of change than the actual records would show. Momentary and even continued increases in muscular tonus occasionally occur during the period of usual tension decrease. Any interpretation for these irregularities would be largely speculative. But the results of our second experiment suggest that these reversals in the general direction of tonus change accompany increased excitation of some other sort. Whether this spread of excitation is caused by the competition of the task idea, with extraneous stimuli, or whether it may result from a pronounced increase in effort, we have no independent means of knowing. It might well be that these momentary increases of tonus are occasioned by the spread of muscular or viceral activity quite irrelevant to the mental work. It must be remembered that our system registers, indiscriminately, the effects of both motor and intellectual activity. The only valid data upon the neural effects of mental work are those regularities which are consistent with careful experimental variation. The rapid rise of tonus coincident with the initiation of an act and its more gradual decline is of this order. We see that this observation, clearly revealed in the tabulated data, holds for both mental and mental-muscular work.

Table 2 also indicates that when a subject is confronted with a series of tasks he tends to develop greater muscular tension in the solution of the first than of the last task. The cause is probably not to be found in fatigue, but rather in the better adjustment of the subject to the later tasks. This is indicated by our data. Task 2, which was reported by all *S*'s as the most difficult, showed the least amount of tension decrease when it occurred late in the experiment. The average decrease for this problem is 1/17 the initial rise when it comes last in the series (B), 1/15 when it comes third (F), 1/5 when second (P) and 1/8 when first (W). Such observations do not seem appropriate to the "fatigue" explanation, where we might, to be sure, expect to find proportionally greater decreases with time. But we are inclined to see in these observations a very real problem in the dynamics of mental work. Some unpublished work of our own has gone to show that decrease of tension during mental work is a function of the amount of initial increase. If this generalization holds for the data now under discussion, the inference is evident. Subjects P and W entered upon the solution of Task 2 in a state of over-preparation or hyper-excitability; whereas the preparation of F and B, who met the same problem later in the series, more nearly approximated the presumptive "energy" requirements of the performance. The further consideration of this problem must await more detailed research.

Our experiments furnish evidence for the existence of two related arhythmic variations in tonus consequent to taking a series of tests: (*a*) a general arhythmical increase at the beginning of the series with gradual diminution in time, and (*b*) a periodic rise and decline of tension coincident with the specific periods of mental work. These changes are even more apparent in the data of the next series.

Series 1b: Discontinuous Operations

The mental activity concerned in a series of reaction trials involves the gross and phasic changes of tonus previously noted. The phasic changes are correlated with the reaction periods, but are frequently superimposed in such a manner as to produce ambiguity. In the majority of cases, however, we found a brief rise and fall to accompany the periods of pre- and post-reaction respectively. The vagaries are interesting. Instead of a post-reaction decline, a rise sometimes occurred. Occasionally a record showed neither an increase or decrease. Inspection of the continuous records indicates a plausible explanation: Reaction trials occurring in fairly rapid succession result in the overlapping of the phase of increased tonus of one trial with the relaxation phase of the preceding. Additional exploratory work has shown that the adequate temporal separation of trials confirms the characteristic features of pre-reaction increase and post-reaction decrease in connection with these momentary acts.

The gross tonus changes also influence the phasic records. The number of trials showing increased tonus in both the pre- and post-reaction periods is greatest near the beginning of the series. The greatest quantitative increases in pre-reaction tension and decreases in post-reaction tension fall within the first half of the trials. The majority of records showing pre- and post-reaction decline of tension come toward the end of the series. The fact that reaction time is also shorter near the end of the series suggests explaining our tonus data as showing the effects of practice or learning. But "practice" and "learning" connote changes in the *products* of mental work, whereas our observations relate to changes in the *agents* involved in the activity. It seems better to refer such changes to organic factors, such as the spread of neural excitation, metabolic variation, or—according to Jacobson's recent proposal—the process of adaptation (13).

Our observations of tonus suggest an explanation of the conflicting results of some of the experiments of the effect of mental work on metabolic rate. The classic Carpenter-Benedict (3) experiments were negative, showing no appreciable increase in metabolism incident to mental work. On the other hand, Becher and Ollson (2) found a positive correlation, as did Laird and Wheeler (15). These last authors report a greater increase after 6 hours sleep than after 8. They surmise that muscular tension was also higher. But they do not discuss the effect this might have on metabolic rate. Dodge reports that Benedict, in commenting upon the positive correlations which were found in some exploratory work, suggested that increased metabolism during metal work might be due to possible tonic accompaniments of the mental activity. Our results apparently confirm Benedict's suggestion.

Table 2*a* summarizes the gross changes in tonus which occur with the discontinuous use of a single complex operation, viz., the compensating manual reaction to seen movement (*Series 1b*). The tonus changes are reported in millimeters of deflection. The first column indicates the individual changes in the level of tension from the end of the period of relaxation to a period approximately one minute after the subject was instructed. (The first ready signal also preceded this reading by 30 seconds.) The second column of the table gives the change in tension from the first trial to a period one minute after the subject is told that the experiment is completed. The actual values are not rendered significant. Equivalent changes in the deflection values for two subjects would not indicate equivalent changes in their quadriceps tonus. These figures show only the general direction and extent of tonus change for individual subjects.

TABLE 2*a*

GROSS TONUS CHANCES IN SERIES 1b

Subject	Tonus change after instruction	Tonus change during reaction series
	(mm.)	(mm.)
P	+13	—13.5
L	+18	—12
Su	+68	—24
F	+26	—30
Se	+37	—15
O	— 7	—23

Table 3 shows the distribution and nature of the phasic tonus records correlated with the reaction periods. Column 1 in this table includes the cases showing an increase in tension during pre-reaction and a decrease during post-reaction (+—). Column 2 includes the cases showing an increase in tension during both pre- and post-reaction periods (++). Column 3 includes the cases of decreased tension in both periods (——). Column 4 includes the cases of no noticeable change (00). The letters *Rt.* after a figure indicate the number of right cases; *Wg.* gives the number wrong, and *A.R.* the number of anticipatory responses. It is interesting to note that all wrong and anticipatory responses resulted in an increased tension in the period of post-reaction.

Series 1c: Alternating Operations

The records of tonus changes occurring with the use of alternative operations confirm the generalizations made under Series 1*a* and 1*b*. The level of tension obtaining for a period of relaxation is noticeably increased with the initiation of mental work of this more difficult

TABLE 3

THE TYPES OF TONIC FLUCTUATION FOUND IN THIRTY REACTION TRIALS

Subject	+— phen	++ phen	—— phen.	00 phen.
P	18 Rt.	2 Rt. 2 Wg. 2 A.R.	4 Rt.	2 Rt.
L	9 Rt.	2 Rt. 2 A.R.	8 Rt.	9 Rt.
Su	21 Rt.	3 Rt. 2 Wg.	2 Rt.	2 Rt.
F	20 Rt.	7 Rt. 1 Wg.	2 Rt.	
Se	21 Rt.	2 Rt. 6 Wg.	1 Rt.	
O	15 Rt. 3 Wg.	4 Rt. 3 Wg.	2 Rt.	3 Rt.
Percentages	60%	21%	10%	.09%

type. The more gradual decrease is also a general rule. Exceptions are worthy of attention since they indicate how very complicated the energetics of mental work may become.

The three graduate students (Se, Sa, and M), who took the psychology tests, gave evidence of very high motivation. These subjects knew that only 10 minutes would be allowed for each test but had no objective indicator of the time. They thus tended to work with great rapidity on the first test given them. Since these tests were completed before the 10-minute period was up, there was a marked tendency for the subjects to work more slowly and carefully on the second test. Time was called on Sa and M before these individuals completed the second test. What happened in the solution of the third test? These two subjects again resumed the more rapid rate and finished before the allotted time! It is interesting to see that this behavior is reflected in the increased tonus changes for the third test. The average 30-second changes are given in Table 4. We have separated the relaxation and examination protocols into two 5-minute periods. The same system of reckoning average change is used here as for Table 2.

The above form of presentation has several disadvantages. Chiefly, it does not show the form of fluctuations in tonus during relaxation and examination. The publication of mean deviations would not aid beyond showing that the greatest variations occur during examination. A pictorial representation of changes in tonus minute by minute is a solution. But the limited use which we should make of this rectorical device does not justify its cost. Instead we shall make our interpretation upon the data as it appears in Table 4. A casual inspection of this table shows that the greatest increase comes in the first half of the examination-period, and the greatest decrease in the last half. Occasionally there is a greater increase than decrease in tension during this last five minutes of examination. These exceptions are marked by an asterisk in the table. Our previous comment on the behavior of Subjects Sa and M in the trials concerned suggests that these changes are correlated with increased effort during the period.

Analogous increases and declines in tension were found in the solution of arithmetic problems by the three naïve subjects. Although the tasks involved the alternative use of several arithmetic operations, the character of the tonus changes is the same as those reported in previous series. Table 5 gives the average 30-second changes in tendon deformation during the solution of each task and its preceding period of relaxation.

We have reserved the consideration of *S*'s reports until the presentation of the tonic changes for Experiment 1 was completed. Practically the same questions were asked in all the series. To these we received a variety of answers.

1. Preparation for a task was frequently reported as following immediately on the reading of instruction. Certain of the more trained subjects reported that the actual realization of a task in the form of a set was not present until some time after instruction.

2. The majority of the *S*'s felt that they worked as hard at the end as at the beginning of the task. One subject remarked that he was "somehow more accustomed to the situation" toward the end. Another said that he came with practice "to know more what to expect."

3. The motivation was usually discovered to favor maximal effort. Desire to excel, hope of reward, and duty toward the experimenter were all mentioned.

4. All three subjects who took the psychological examinations reported a tendency to think about their answers to the preceding test. With simpler work, the tendency was to dismiss all thought of the preceding problem very soon. Anticipation was more usually reported for the first task of a series than for later ones. We are not able to say that such residual or anticipatory effects find a reflection in the records of relaxation tension. There may well be such a reflection. Both possible adjustments to irrelevant internal stimuli make the relaxation changes ambiguous. In one pronounced case (*Series 1c-W, 3*) we do know that the subject became very irritated

TABLE 4

AVERAGE CHANGES IN TONUS DURING EXAMINATIONS IN PSYCHOLOGY

		Test 1				Test 2				Test 3			
S	T.O.	Relax. changes		Exam. changes		Relax. changes		Exam. changes		Relax. changes		Exam. changes	
		1-5m	6-10m	1-5m	6-10m	1-5m	6-10m	1-5m	6-10m	1-5m	6-10m	1-5m	6-10m
Sa.	1,2,3	—1.3	— .6	+ 6.2	—4.0	+.2	+5.0	+ 3.2	—5.4	+2.1	+8.0	+ 8.4	+2.5*
M	2,3,1	+4.3	+3.1	+15.5	+1.3	—.5	— .3	+10.3	—6.8	—2.0	— .1	+11.2	—3.8
Se	3,1,2	— .3	—1.2	+ 3.5	— .2	+.8	+ .2	+ 2.3	—1.5	— .3	— .5	+ 2.1	—4.6
Av.		+ .9	+1.3	+ 8.75	— .97	+.17	— .83	+ 5.27	—4.57	— .07	+2.47	+ 7.23	—1.97

TABLE 5

AVERAGE TONUS CHANGES IN THE SOLUTION OF ARITHMETIC PROBLEMS

		Problem 1				Problem 2				Problem 3			
S	T.O.	Relax.		Work		Relax.		Work		Relax.		Work	
		1-5m	6-10m	1-5m	6-10m	1-5m	6-10m	1-5m	6-10m	1-5m	6-10m	1-5m	6-10m
R	1,2,3	+3.8	+2.2	+12.3	— .2	+3.1	+4.6	+14.6	—2.2	+ .3	—2.1	+3.2	—4.0
W	2,3,1	— .3	— .2	+ 5.2	— .1	+3.0	+ .5	— .2	—1.1	—3.0	+ .5	—2.0	+ .2
F	3,1,2	+ .8	— .5	+ 8.9	—4.6	+ .3	—1.6	+ 8.5	— .7	+4.9	+2.3	+5.3	—1.2
Av.		+1.43	+ .5	+ 8.8	+1.4	+2.13	+1.17	+ 7.63	—1.33	+1.63	+2.33	+2.03	—1.66

at the delay, feeling that it was time he met another appointment. His tonus record showed a steady rise during the period of the usual relaxation phase, with a quick *drop* when the next task was announced. An instance of this sort shows how susceptible our tonus records are to all neural excitations, and what experimental caution is therefore necessary in drawing valid inferences.

EXPERIMENT 2

We have seen from Experiments 1*a,* 1*b,* and 1*c,* that the initiation of a task causes increased tonus which normally decreases during the solution of the task. Experiment 2 considers the question of what occurs in terms of tonus changes when the solution of a task is interrupted or disturbed. Decisive results were obtained by timing our interruptions so that they usually fell in the period of release from initial muscular tension. In such cases the direction of change was usually reversed by interruption.

The critical tonus-changes in this experiment are those occurring during equivalent periods of interrupted and uninterrupted mental work. Although the level of tonus was taken every 30 seconds for the first three minutes of work, only the average 30-second change during the critical periods of interruption (second minute) are given in the tables below. The data are separated according to type of work interrupted. Table 6 summarizes the results obtained on purely mental tasks and Table 7 serves for mental-muscular tasks. The rhetorical purposes which these tables serve has resulted in a scattering of the records of individual subjects. Inclusion of the subjects' names would have complicated the tables without increasing clarity. The important fact is the striking difference in the direction of tonus change for equivalent periods of interrupted and uninterrupted work. The actual amount of change varied, of course, from task to task and from subject to subject. The factors determining the amount of increase or decrease include the nature of the task, the character of the interruption, and the motivation of the subject. While these actual changes are not expressed in terms of a common denominator, we can average them for purposes of comparison. This is permissible because each record for interrupted work is paralleled by one taken on the same subject during a related period of uninterrupted labor.

Usually our subjects continued their task notwithstanding inter-

TABLE 6
Tonus Changes during Equivalent Periods of Interrupted and Uninterrupted Mental Work

	Work interrupted by					Uninterrupted work					Averages for	
Tasks	C₁	C₂	C₃	D₁	D₂						I.W.	U.W.
1	+12.5	+15.0	— 2.0	— 1.5	+12.0	+ 1.0	+ 1.5	— 2.5	— 6.5	— 4.5	+ 7.2	— 3.2
2	+ 1.0	+ 5.0	+ 5.0	— 3.2	+ 2.5	— .5	— 6.0	+ 3.0	—12.0	— 5.1	+ 2.1	— 4.1
3	+17.0	+ 1.0	— 7.0	+25.0	+ 3.0	— 4.0	— 6.5	— 3.5	— 2.0	—16.0	+ 7.8	— 6.4
4	+ 2.0	+20.0	+ 1.0	+ 1.2	— .5	—22.5	—20.0	— 1.5	— 3.0	— 2.0	+ 4.7	— 9.8
5	+21.0	+ 0.0	+ 5.5	5.0	+ 9.2	— 2.1	— 2.0	— 3.5	— 1.2	— 3.0	+ 8.1	— 2.3
6	+11.0	+ 8.0	+ 5.0	+ 3.5	— 4.0	— 2.5	— 3.5	+ 0.0	— .5	+ 3.0	+ 4.7	— .7
7	— .2	0.0	+19.0	+ 1.0	+ 8.0	—19.0	+ 5.0	—12.0	— 5.0	— 8.0	+ 5.5	— 7.8
8	+ 4.5	+ 1.0	— 2.0	+ .5	+ 1.5	—35.0	— 6.5	—22.5	— 3.0	—10.0	+ 1.1	—15.4
9	+ 4.0	+ 8.2	+ 7.5	+27.0	+ 3.5	— 7.0	— 3.5	— 2.0	+ 8.0	—12.2	+10.0	— 3.5
10	— 2.0	+28.5	+ 3.0	+ 2.0	+12.0	— .4	—12.0	+ 7.0	— 1.0	+ 1.0	+ 8.7	— 1.0
	Average change +6.00					Average change —5.32					+ 5.99	— 5.41

TABLE 7
Tonus Changes during Equivalent Periods of Interrupted and Uninterrupted Mental-Muscular Work

	Work interrupted by					Uninterrupted work					Averages for	
Tasks	C₄	C₅	D₃	D₄	D₅						I.W.	U.W.
1	+ 1.5	+ 8.0	— 3.0	— 2.0	+10.0	— .5	— .5	+ 4.5	— 8.0	— 6.0	+ 2.9	— 2.1
2	— 4.0	+ 3.5	+ 5.0	+ 3.5	+ 1.5*	—2.0	— 4.5	+ 2.1	—25.5	— 1.0	+ 1.9	— 6.2
3	+ 1.0	+ 1.0	+ 5.5	+ 1.5	+ 4.0	— .5	— 8.0	— 3.0	— .5	+ .5	+ 2.6	— 2.3
4	+18.0*	+ 3.0	+12.0	+ 1.5	—10.0	+3.0	+ 6.5	— 5.5	— 2.5	—20.5	+ 4.9	— 4.0
5	+ 5.0	+ .2	+ 2.5	+ 1.2	+ 3.0	—3.0	+ 1.2	— 7.5	—16.0	— 4.0	+ 2.4	— 5.8
6	+12.5	+ 4.5	+ 6.5	+ 3.0	+ 0.0	+4.1	—13.5	— 2.5	—10.5	— .5	+ 5.3	— 4.6
7	+ .5	— 2.0	+ 0.0	+ 1.5	— 1.5	—2.0	—25.0	— 2.5	+ 3.0	+ 1.5	+ .3	— 5.0
8	+ 0.0	+ 5.0	+10.0	+ 3.0	+10.5*	—8.0	— 4.5	+ 1.5	— 1.0	—12.2	+ 5.7	— 4.8
9	+ 6.5	+14.0	+ 3.0	+ 8.0	— 2.5*	+3.5	+ .7	+ 5.0	—30.0	— 0.0	+ 5.8	— 5.6
10	— .2	+ .5	+22.5	+12.0	+ 6.0	—1.7	—10.0	— 8.0	— 0.0	— 2.0	+ 8.1	— 4.3
	Average change +3.98					Average change —4.47					+3.99	—4.47

ruption. In the case of more serious disturbances actual work was halted until the disturbance was removed. There are only four cases where the subject did not resume his task. These cases are marked in the tables by asterisks; three of them occur in connection with in-interruption D_3, where the subject is *told* to begin another task. The other case cannot be counted properly (Task 4, C_4) because the fuse blew.

A comparison of average tonus changes for equivalent periods of interrupted and continued mental work indicates an increased spread of neural excitation when interruption is attempted. The increased spread to centers of tonic activity probably serves to reinforce the original task set and to inhibit responses to the irrelevant stimuli. Such an explanation is in agreement with the observations on the effect of distractors on attention. It is also interesting to note that Jacobson has recently reported that the progressive relaxation of a subject is interrupted when the subject is disturbed.

While there are no noticable differences in the effect of accidental (A) and deliberate (D) interruptions on the level of tonus, the clearest results were probably obtained with the deliberate interruptions. Very little tension accompanied the adjustment of subjects W and R to the presentation of these "interruptions" during a relaxation period. But such an accidental interruption as A_1 (the falling curtain) resulted in a pronounced increase in tension when given during relaxation. This indicates that the increased spread of excitation accompanying task interruption is not due entirely to the attempt of the subject to continue work. Frequently a part of the spread is due to an emotional adjustment to the new stimulating conditions. The deliberate interruptions (*D.I.*) would produce relatively little spread except as they present a complication to the task-situation.

Discussion

Our results have certain bearings on the neural-muscular consequences of completed and interrupted mental work. In Experiment 1 we found that beginning a task was accompanied by an increased tension in a remote muscle-group. This tension decreased as the work progressed towards completion. In Series 1*c* we found some reason for believing that the amount of initial tension corresponded to the degree of effort. Apparently there is also a relationship between the diminution of this tension and adequacy of the reaction.

Experiment 2 showed that continuation of a task during distracting episodes of interruptions involved a further increase in muscular tension. These observations may be taken as presumptive indications of the action of the entire neural mechanism. The initial spread of excitation to centers producing tonic activity is a consequence of a "set" which is initiated by instruction. This set is normally ended by completion of the task. The neural mechanism thus returns to a state of equilibrium which approximates its *status quo ante*. Interruption of a task introduces a complication into this normal progress towards equilibrium. The accompanying increase in muscular tension cannot be dismissed solely as the result of spread from a new excitation. Instead there is evidence of a reciprocal effect, the tonic activity apparently reinforcing the neural action involved in the task-set.[2] We shall return to the question of reinforcement later.

The considerations outlined above re-open the important question suggested by Lewin's recent experiments on interrupted work. Why does the organism tend to continue a task notwithstanding interruption? It will be recalled that Lewin places the energetics of this behavior in a "psychic tension" specific to the interrupted task. This tension is alleged to exert a stress which is responsible for task resumption. Completion of a task is the most common way by which tensions are discharged and equilibrium restored. In short, according to Lewin, all behavior may be viewed as the result of specific psychic tensions discharging in specific fields.

We shall not attempt a theoretical criticism of Lewin's doctrine, though there are internal difficulties. Our objections to its application to the question of task resumption are (*a*) that it is opposed to neural tradition and (*b*) that a simpler explanation in terms of that tradition is possible.

(*a*) *Opposition to Neural Tradition.* The concept of "psychic tension" which is advanced to explain the energetics behind resumption of interrupted acts must be understood in terms of hypothetical energy-transformations. Viewed in this light, the concept appears to fit best with the idea of some special form of "mental" energy. It is probable, however, that the parallelistic postulates of *Gestalttheorie* would commit Lewin to the assumption that his tensions are correlated with nervous stresses. This assumption would seem to embrace a quantum theory of nervous energy, which is open to serious objection by physiologists (1) and psychologists[3] who favor physiological interpretation. The energy equivalent of a "psychic tension" must be capable of being stored as a specifically directed quantum to be released at a later time. In no other way can we understand how the kinetic energy of a tension-producing present behavior may, upon interruption of that behavior, become potential. Resolves, wants, etc., are alleged to be manifestations of these unexpended tensions which are stored up and discharged appropriately. Aside from its attractive and plausible explanation of these mental experiences, the doctrine of psychic tension has little to support it. The energy transformations which it is forced to assume are opposed by the results of physiological and psychophysiological research.

---

[2] We showed that when the same disruptive stimulus was presented during relaxation a smaller increase in tension occurred than with its presentation during task-solution. More direct evidence of a reciprocal effect is to be found in (4).

(*b*) *A Simpler Explanation.* There remains the possibility of interpreting the resumption of interrupted acts in terms of neural-muscular research. Our data are in agreement with those of Lewin. Initiation of a task does set up an organic disturbance which is normally equilibrated by completion. Tasks once initiated tend to be continued in spite of interruption. The experimental facts indicate that these interrupted task-sets never completely lose their directive influence on behavior. Anything beyond this is an assumption.

We are in a somewhat better experimental position than Lewin, however, when it comes to a discussion of the underlying processes which condition this behavior. We measured the neural-muscular effects of disturbing interruptions. Our interpretation is concerned particularly with the relation of these effects to the total behavior.

We assume that when the organism is under the dominance of a particular task-set, distracting stimuli may facilitate as well as inhibit the objective performance. This is probable from the literature on "distractors" (16). Our tension records suggest that tasks are continued, notwithstanding distraction, by virtue of some *reinforcement* to the task-set either derived from or expressed by the tonic activity of the bodily musculature. Instead of assuming the directive influence of a blocked "psychic tension," we find a neural analogy of task resumption in the *competition* of rival impulses for a final common path. While this analogy is far too simple to be exact, experiment has demonstrated that the principle still holds for the behavior of the intact human nervous system.[4] In the case of task resumption, the competition is rather between two unrelated response patterns for directive control of behavior. In this competition the pattern connected with task solution has this great advantage over all others conditioned by rival stimuli; it is usually in full operation at the time when these other patterns are sub-liminal or in process of development. Because it appears more significant, it consequently may be reinforced through the action of other centers. There is no assumption that the action of these centers is itself selective. Whichever response-pattern is dominant may obtain reinforcement from their activity.

---

[3] See the discussion by Dodge on the quantum theory of neural energy in (8, pp. 114-116).

The continuation of interrupted mental work involves not only reinforcement of the dominant response pattern, but also the reciprocal *inhibition* of irrelevant responses. Frequently the balance of inhibition and reinforcement favors the dominance of these other patterns. In this case, task resumption is conditioned upon the ability of the competing task-set (through summation or otherwise) to regain directive control. In this struggle inhibition may take a form opposed to the interrupting response. There are many instances of this sort in both Lewin's and our own work; for example, the case where the subject moves his arm away when the experimenter reaches over to appropriate his pencil.

In only one case would it seem that neural competition, reinforcement, and inhibition fail to explain the resumption of interrupted work; and here these principles are not needed. It may be claimed by some that the request to do something unrelated to the task is actually carried out, and that task resumption, at a much later time, is unrelated to the original competition between task-set and interruptive response. But here we may simply assume that the later occasion *itself* serves to reinstate the original task-set.

Our explanation applies specifically to the resumption of interrupted acts. But a related explanation may be possible for Lewin's other observations in the field of action. We shall not attempt to extend our interpretation. It is obvious that this project must await psychophysiological methods adequate to the repetition of these other problems. We believe, however, that our discussion indicates the direction which such a reinterpretation of Lewin's data may take. Competition, reinforcement, and inhibition of neural impulses are well known physiological principles. Although they were discovered in the reflexes, there is no a priori reason why they cannot be applied to the dynamics of mental work. In fact, much careful experimentation indicates that such principles are needed to explain the behavior of the intact human organism. The resort to such concepts as "psychic tension" is, on the other hand, a hypothesis without support in the neurological tradition of energy-transformations.

---

[4] This process has been demonstrated in the intact human by Dodge (7).

Summary

Mental work involves an initial increase in muscular tension which decreases as the performance progresses towards completion. Experiments were conducted on 10 subjects and with tasks which utilized (*a*) a continuous unitary operation, (*b*) a discontinuous unitary operation, and (*c*) several alternative operations. Other experiments studied the effect of interruption on the tonus accompaniments of mental work. Twenty tasks (10 mental and 10 mental-muscular) were given to 10 subjects. Half of these tasks were interrupted one minute after they were begun. About 98% of the interrupted tasks were either continued during interruptive disturbance or resumed one minute after the interruption began. A comparison of the tonus change during equivalent periods of interrupted and uninterrupted mental work showed a regular and notable increase in tension during the periods of interruption. The results of this paper form the basis for the argument that the dynamics of completed and interrupted mental work may be understood in neurophysiological terms. Instead of the doctrine of "psychic tension" as proposed by Lewin to account for these facts, only such concepts have been utilized as competition, reinforcement, and inhibition. The suggestion is made that eventually all of Lewin's important research in the field of action may be reinterpreted in these or related neural terms.

References

1. Adrian, E. D. The conception of nervous and mental energy. *Brit. J. Psychol.*, 1923, **14**, 121-125.
2. Becker, F. C., & Ollson, O. Metabolism during mental work. *Skand. Arch. f. Physiol.*, 1914, **31**, 80-197.
3. Benedict, F. G., & Carpenter, T. M. The influence of muscular and mental work on metabolism. *U. S. Dept. Agric., Bull.*, 1909, No. 209.

4. Bills, A. G. The influence of muscular tension on the efficiency of mental work. *Amer. J. Psychol.,* 1927, **38**, 227-251.
5. Brown, J. F. The methods of Professor Lewin in the psychology of action and affection. *Psychol. Rev.,* 1929, **36**, 200-221.
6. Dodge, R. Mental work. *Psychol. Rev.,* 1913, **20**, 1-42.
7. ———. Adequacy of reflex compensatory eye-movements, including the effects of neural rivalry and competition. *J. Exper. Psychol.,* 1923, **6**, 177-181.
8. ———. Theories of inhibition. Part 1. *Psychol. Rev.,* 1926, **33**, 106-122.
9. Freeman, G. L. The measurement of tonus by the method of tendon deformation. *Amer. J. Psychol.,* 1930, **42**, 581-592.
10. Golla, F. L. The objective study of neuroses. *Lancet,* 1921, **2**, 115-122; 215-221; 265-270; 373-379.
11. Gopalaswami, M. V. A note on the correlation between the psychogalvanic reflex and learning effort. *Indian J. Psychol.,* 1926, **1**, 35-38.
12. Ilzholfer, H. Ueber den Einfluss der Trainings auf Grundumsatz und Arbeitsnutzeffekt. *Arch. f. Hygiene,* 1924, **94**, 317-328.
13. Jacobson, E. Differential relaxation during reading, writing and other activities as tested by the knee jerk. *Amer. J. Physiol.,* 1928, **86**, 675-693.
14. ———. Progressive relaxation. Chicago: Univ. Chicago Press, 1929. Pp. 379.
15. Laird, D. A., & Wheeler, W. What it costs to lose sleep. *Indus. Psychol.,* 1926, **1**, 694-696.
16. Morgan, J. J. B. The overcoming of distraction and other resistances. *Arch. Psychol.,* 1916, **35**. Pp. 84.
17. Ovisankina, —, & Lewin, K. (In press.) See (5).
18. Travis, R. C. A study in the measurement of muscle tonus and its relation to fatigue. *J. Exper. Psychol.,* 1924, **7**, 201-212.
19. Tuttle, W. W. The effect of attention or mental activity on the patellar tendon reflex. *J. Exper. Psychol.,* 1924, **7**, 401-419.

Yale University
New Haven, Connecticut

LES CHANGEMENTS DE TONUS PENDANT LE TRAVAIL MENTAL COMPLET ET INTERROMPU

(Résumé)

Dans le travail mental il s'agit d'une croissance initiale de la tension musculaire qui tend à décroître comme le travail approache sa termination. On a fait des expériences avec dix sujets et des tâches utilisant (1) une opération unitaire continue, (2) une opération unitaire non continue et (3) plusieurs opérations alternatives. D'autres expériences ont étudié l'effet de l'interruption sur les accompagnements de tonus du travail mental. On a donné vingt tâches (10 mentales et 10 mentales-musculaires) à 10 sujets. On a interrompu la moitié de ces tâches une minute après leur commence-

ment. Environ 98% des tâches interrompues ont été ou continuées pendant l'interruption ou reprises en une minute après le commencement de l'interruption. Une comparaison des changements de tonus pendant des périodes équivalentes de travail mental interrompu et non interrompu a montré un accroissement régulier et frappant de tension pendant les périodes d'interruption. Les résultats de cette étude forment la base du raisonnement que l'on peut comprendre la dynamique du travail mental complété et interrompu en termes neuro-physiologiques. Au lieu de la doctrine de "tension psychique," comme proposée par Lewin pour rendre compte de ces faits, on a employé seulement des conceptes tels que compétition, renforcement et inhibition. On offre la suggestion que l'on pourra à la fin interpreter de nouveau toutes les recherches importantes de Lewin en ces termes nervaux ou en termes semblables.

FREEMAN

VERÄNDERUNGEN IM TONUS WÄHREND BEENDETER UND UNTERBROCHENER GEISTIGER ARBEIT

(Referat)

Geistige Arbeit schliesst eine Erhöhung der Muskelspannung ein, die nachzulassen pflegt, wenn die Verrichtung gegen die Beendigung vorrückt. Es wurden Experimente an 10 Versuchspersonen vorgenommen und zwar mit Aufgaben, welche a) einen kontinuierlichen einheitlichen Vorgang, b) einen unterbrochenen einheitlichen Vorgang und c) mehrere alterierende Vorgänge gebrauchen. Andere Experimente untersuchten den Effekt der Unterbrechung bezüglich des Tonus, der die geistige Arbeit begleitet. Zwanzig Aufgaben (10 geistige und 10 geistig-muskuläre) wurden 10 verschiedenen Versuchspersonen gegeben. Die Hälfte dieser Aufgaben wurden eine Minute nach dem Anfang unterbrochen. Etwa 98% der unterbrochenen Aufgaben wurden entweder wahrend der Unterbrechung fortgesetzt, oder eine Minute nach der Unterbrechung wieder aufgenommen. Ein Vergleich der Tonusänderungen während equivalenter Zeitabschnitte unterbrochener und ununterbrochener geistiger Arbeit weist eine regelmässige und merkbare Erhöhung der Spannung während der Unterbrechungszeit auf. Die Ergebnisse dieser Untersuchung bilden die Grundlage für die Behauptung, dass die Dynamik vollendeter und unterbrochener geistiger Arbeit neuro-physiologisch verstanden werden kann. Statt der Lehre von der "psychischen Spannung," wie sie von Lewin zur Erklärung dieser Tatsachen vorgeschlagen wurde, wurden nur Begriffe wie Wettkampf (competition), Verstärkung (reinforcement) und Hemmung (inhibition) verwendet. Man schlägt vor, dass vielleicht später alle von Lewins Untersuchungen im Gebiet der Tätigkeit in diese oder eine ihr verwandte Nerventerminologie umgesetzt werden können.

FREEMAN

SHORT ARTICLES AND NOTES

EDITED BY
F. C. BARTLETT
V. BOROVSKI
F. KIESOW
MATATARO MATSUMOTO
WALTER MILES
R. W. WILCOCKS

NOTICE OF SPECIAL VOLUME

CARL MURCHISON

Along with most of the other psychological journals, we have been facing for some time the problem of diminishing the large number of manuscripts which have been accumulating steadily in our files and which have resulted in lengthening unduly the amount of time necessary to get a manuscript published after it has been received from its author. This problem has been met in various ways by some of the other psychological journals, and we have finally decided on a method which to us seems wisest under the circumstances.

We have decided to publish an entire volume in one issue, this issue to appear between two current volumes. This method will result in the saving of an entire year and will clear our files to within approximately six months of publication. We have decided on this method chiefly because it adds no additional financial burdens to the author and extends to all authors alike equal benefits in the reduction of the length of publication time.

This volume, Volume 4, will not be sent out on subscription, nor will it be involved in the Club Rate plan. It must be ordered directly by each subscriber at the regular subscription price of $7. Volume 5 will be issued as the current volume during the year 1931.

THE DEVELOPMENT OF HALF-CENTERS IN RELATION TO THE QUESTION OF ANTAGONISM IN REFLEXES[1]

G. E. COGHILL

Graham Brown (2) has suggested "that the spinal reflex is of the nature of a distorted 'beat' of the rhythmic progressive phenomena," and he has stated it as an inference "that the rhythm is conditioned by properties of the mutual interconnexions of the 'half-centres,' and that it occurs during their equal activation." The activation, he thinks, may come from exteroceptive or proprioceptive sources, or from endogenous conditions after the manner of the respiratory rhythm.

---

[1] Read before the Ninth International Congress of Psychology.

"Half-centres," according to Graham Brown, are units of function that are linked together in pairs to form the center as a whole that controls a particular movement, such as flexion or extension. The two "half-centers" of a pair are regarded as mutually antagonistic and as requiring appropriate excitation and inhibition in order to produce the distorted "beat," which is a reflex. It is this type of organization and function that I wish to discuss upon the basis of my studies of correlated structure and function in the development of Amblystoma (5).

In the early stage of motility in Amblystoma, the motor mechanism consists of paired cell columns which extend through the medulla oblongata and spinal cord, and which, in the anterior region, are interconnected by commissural neurons (4).

The effectors that are innervated by these two motor columns are the paired series of muscle segments. There are no motile appendages. The organization of the nervous system is such that stimuli enter the anterior part of the columns, and conduction is from the head region tailward (3). The muscular response is for that reason a wave of contraction that is propagated from the head tailward.

The contractions are at first restricted to the anterior part of the body, and are very slow. As the animal advances in development, the wave of contraction becomes longer, and its speed increases. The contractions may be to the right or to the left, but in the earlier phase of development they are not at the same time in different directions in different levels. Later, however, the first contraction to the right, for example, is followed by a contraction to the left before the first contraction has passed entirely through the trunk. A rhythmic series of contractions of this kind eventually effects locomotion.

We see here, accordingly, rhythmic action of the motor centers as the basis of "progressive phenomena," as Graham Brown expresses it; and we find very definite paired "half-centers" as the nervous mechanism that excites these rhythmic phenomena.

In the development of motility, according to this plan, there is inherent possibility of mutual antagonism between the paired effectors. But antagonism of the kind that produces rigidity in reflexes under experimental conditions has not been observed. The movements nearly always appear to be appropriately coordinated. A possible exception occurs occasionally about the time swimming begins, when the rhythmic contractions are of less amplitude and more rapid than usual. Nevertheless, such movements effect locomotion.

Locomotion, accordingly, is developed by the expansion and speeding up of the original simple flexure which becomes rhythmic; and in what may be regarded as normal cases, locomotion develops without evidence of antagonism between the potentially antagonistic paired effectors. The explanation for the lack of antagonism in this rhythmic system is found in the development of the neuromotor mechanism that activates it.

The center of control of these movements begins as a nervous matrix of relatively uniform structure that spreads across the median plane of the central nervous system ventrally. About the time motility begins this originally unitary center divides longitudinally into two lateral halves, each of which comes to be in direct control of the effector of its own side. As the contractions increase in amplitude and speed, and become rhythmic, collaterals grow from the motor neurons of one "half-center" into a synaptic relation with commissural cells which terminate in the "half-center" of the other side. These collaterals are presumably inhibitory, and as such prevent the antagonism between the two "half-centers." Their action would also facilitate the speeding up of the contraction wave.

In Amblystoma, then, the paired "half-centers" of progression movement emerge from a unitary matrix of motor neuroblasts. According to observations by one of my associates, Doctor Angulo y Gonzalez (1), the several columns of motor cells in the ventral horn of the spinal cord of the albino rat emerge in the same manner from a unitary cell mass, and their emergence is correlated with the development of movement in the trunk and limbs. Until contradictory data are at hand, therefore, the development of "half-centers" in relation to motility in Amblystoma may be regarded as typical for vertebrates generally.

In this physiologic and anatomic development there operate two principles that are basic to the understanding of the antagonism of reflexes; first, "half-centers" emerge within an originally unitary and integrated system, and secondly, reflexes are facilitated by the growth of neurons after they begin to function as conductors.

REFERENCES

1 ANGULO Y GONZALEZ, A. W. Neurological interpretation of fetal behavior: the motor-cell columns of the albino rat before birth. *Anat. Rec.,* 1929, **42**, 17.

2. BROWN, T. G. On the nature of the fundamental activity of the nervous centres; together with an analysis of the conditioning of rhythmic activity in progression, and a theory of the evolution of function in the nervous system. *J. Physiol.,* 1914, **48**, 18-46.

3. COGHILL, G. E. Correlated anatomical and physiological studies of the growth of the nervous system of amphibia. III. The floor plate of *Amblystoma. J. Comp. Neur.,* 1924, **37**, 37-70.

4. ————. Correlated anatomical and physiological studies of the growth of the nervous system of amphibia. V. The growth of the pattern of the motor mechanism of *Amblystoma punctatum. J. Comp. Neur.,* 1926, **40**, 47-96.

5. ————. Anatomy and the problem of behavior. (Lectures delivered at University College, London.) London: Cambridge Univ. Press; New York: Macmillan, 1929. Pp. xii+113.

The Wistar Institute of Anatomy and Biology
Philadelphia, Pennsylvania

MAZE PERFORMANCE A FUNCTION OF MOTIVATION AND OF REWARD AS WELL AS OF KNOWLEDGE OF THE MAZE PATHS[1]

EDWARD CHACE TOLMAN

The classical account of maze performance is too simple. It conceives the rat as responding to a mere series of moment-to-moment stimuli. It is not adequate to care for the results of recent experiments in which motivation and reward are themselves varied.

I shall briefly review some of these latter results and then suggest what seems to me a more adequate formula.

Simmons (8) ran different groups of rats under different combinations of motive and goal-object. She found that some of these combinations caused more rapid learning—that is, a sooner perfect performance—than did others.

Szymanski (9) kept feeding-conditions constant but varied degree of hunger. He found that rats when not hungry would run incorrectly even though when previously hungry they had made a perfect performance.

Hamilton (7) varied the conditions of feeding. She found that a delay inserted between the rat's reaching of the goal-box and its receiving of food retarded learning.

Washburn (12) divided her mice, on a basis of general performance in running and readiness to eat, into a hunger-driven group and an activity-driven group. She found that the hunger-driven group were better at avoiding blinds than were the merely activity-driven group.

The other experiments to be reported have been performed at California but have not yet, for the most part, appeared in the literature.

Tolman and Robinson and Honzik (11) compared the performance of very hungry rats with that of less hungry rats. They found that the less hungry rats seemed to have a preference for long blinds, whereas the very hungry rats soon avoided these longer blinds.

Blodgett (1) ran hungry rats under non-reward conditions. They showed little apparent learning. He then, however, suddenly rewarded them, and their error and time curves immediately dropped. They gave evidence, that is, of a latent learning which manifested itself when the reward was introduced.

---

[1] Read before the 9th International Congress of Psychology. Yale University, New Haven, Conn. Sept. 6, 1929.

Tolman and Honzik (10) started their rats with reward, and then, when they had practically learned, took this reward away. The error and time curves went up. In fact the animals did worse than they would have if they had never had food at all. They seemed to be hunting for the lost food.

Bruce (2), in a similar experiment, tried the effect of increasing hunger during the non-reward period. He found that the hungrier the rats, the poorer their performance—the more, that is, they seemed to be hunting for the lost food.

Elliott (3) changed rats part way through learning from bran mash to sunflower seed. Their errors and times went up. In fact, they did more poorly than a control group which had had sunflower seed from the beginning. They seemed to be hunting for the lost mash.

Elliott (4) changed rats from the combination of thirst, with water as goal-object, to that of hunger, with food as goal-object. On the first day of the change, when, that is, the rats were first hungry and before they had yet found food, their errors and times went up. After having once found food, however, their performance dropped to its former level.

Elliott (5), in another experiment, changed three groups from food as goal to water as goal. The first group were both very hungry and very thirsty. The change from food to water in no way showed in their performance. The second group were very hungry and only slightly thirsty. Upon being changed from food to water they shot up in both errors and times. The third group were slightly hungry and very thirsty. Upon being changed to water, they went down in errors and in times.

Finally, Mrs. Williams (13) tried the possibility of a "conditioned" goal-object. She first taught her rats to find food in a white-black discrimination box. She then ran these same rats in a maze and, part way through learning, introduced this discrimination box, *but with no food in it,* as the goal for the maze. Upon first finding it at the end of the maze, these rats (who till then had received no reward in the maze) improved their maze performance. After a few more trials—that is, when these rats discovered that here at the end of the maze this box no longer "meant" food—their maze performance reverted to its previous poor level.

So much for the experiments.

What new categories for the description and prediction of maze performance do they suggest? I propose the following. And here I must acknowledge my fundamental indebtedness to Dr. Elliott whose own discussion (6) I have taken over and elaborated.

Maze-performance, Elliott would declare to be describable as a function of three main independent variables:

1) *Knowledge of the maze paths.* This is to be conceived as a deposit left primarily as a result of exploratory runs. In the present context it can be considered as practically proportional to the number of such runs.

2) *Motivation.* This is the physiological state, such as hunger, thirst, sex-excitement, pain, or what not, which makes the animal active and exploratory.

3) *The goal-object actually provided at the exit.* This needs no further descriptive comment.

In the experiments just described it was primarily variations in 2 and 3, that is, in motivations and goal-objects, which were investigated. But it appears now, further, that in order completely to explain the results a fourth intermediate variable, resulting from this motivation and this actually provided goal-object, and also from certain auxiliary conditions, must be introduced. This fourth variable may be designated:

4) *The reward-value of the provided goal-object.*

And the total conditions determinative of this reward-value seem to me to be the following four: (It is here only that I have elaborated upon Dr. Elliott's account.)

a) The actual physiological appropriateness of the provided goal-object, given the motivation.

b) A seeming appropriateness in the goal-object, given this motivation and given certain specific past experiences of the rat.

c) The degree to which the provided goal-object is expected.

d) The degree to which this goal-object may happen to be a substitute for some previously provided goal-object.

Each of these may be explained further:

a) *The goal-object's actual physiological appropriateness.* The introduction of this concept is required by such facts as that food at the exit will induce a good performance, if hunger be the motive, and a poor performance, if thirst be the motive, and so on. And it is also required by such facts as those presented by Miss Washburn and by Tolman and Robinson and Honzik that, if the motivation be one of activity or slight hunger only, performance (defined in terms of getting to the exit) will be poor—or in other words, the continued entering of blinds, especially of long blinds, may be favored.

b) *The goal-object's seeming appropriateness, given the motivation and given the rat's specific past experiences.* The introduction of this concept is required by such findings as that of Mrs. Williams, that a goal-object not actually appropriate—for example, an empty discrimination-box—may, none the less, because of specific past experiences, *seem* appropriate and have therefore, for a time, a high reward-value.

c) *The degree to which the given goal-object is expected.* This concept is required to allow for the obvious fact (appearing in all the experiments in which a shift in the goal-object was made during the course of the experiment) that only after the new goal-object had been experienced one or more times did it induce a change in maze performance.

d) The degree to which the given provided goal-object may happen to be a substitute for some previously provided goal-object. This last concept is required by the fact that a new goal-object, such as sunflower seed (as demonstrated by Elliott) or no goal-object (as demonstrated by Tolman and Honzik and by Bruce) will have less reward-value when introduced part way through learning than it would have had if present from the beginning.

One last word. I am aware that the introduction of such anthropomorphic concepts is offensive. I apologize. And, further, I am willing to hope that perhaps some day they may be abandoned in favor of very detailed lists of underlying stimulus-response bonds or in favor of some picture of underlying configurational stresses, strains and tensions—working themselves out according to the laws of "Closure," *"Prägnanz,"* "Least Action," and the rest. For the present, however, I, for my part, must declare that, *when it comes to rats,* neither the system of Thorndikian bonds, nor that of Gestaltian configurations seems to me, as yet, sufficiently detailed or sufficiently specific to be as fruitful for further research as are the above anthropomorphisms.[2]

REFERENCES

1. BLODGETT, H. C. The effect of the introduction of reward upon the maze performance of rats. *Univ. Calif. Publ. Psychol.,* 1929, **4**, 113-134.
2. BRUCE, R. H. The effect of removal of reward on the maze performance of rats. (M.A. thesis.) University of California, 1929.
3. ELLIOTT, M. H. The effect of change of reward on the maze performance of rats. *Univ. Calif. Publ. Psychol.,* 1928, **4**, 19-30.
4. ————. The effect of change of "drive" on maze performance. *Univ. Calif. Publ. Psychol.,* 1929, **4**, 185-186.
5. ————. The effect of appropriateness of reward and of complex incentives on maze performance. *Univ. Calif. Publ. Psychol.,* 1929, **4**, 91-98.
6. ————. "Motivation and reward in relation to maze performance." (Ph.D. thesis.) University of California, 1929.
7. HAMILTON, E. L. The effect of delayed incentive on the hunger drive in the white rat. *Genet. Psychol. Monog.,* 1929, **5**, 131-207.
8. SIMMONS, R. The relative effectiveness of certain incentives in animal learning. *Comp. Psychol. Monog.,* 1924, **2**, No. 7. Pp. 79.
9. SZYMANSKI, J. S. Ein experimenteller Beitrag zur Analyse der bei Entstehung neuer gewohnheiten mitwirkenden Faktoren. *Pflüg. Arch. f. d. ges. Physiol.,* 1918, **170**, 197-220.

---

[2] This last sentence was written before I had heard Dr. Kurt Lewin's stimulating presentation of his doctrine of *Feldkräfte, Topologie,* and *Aufforderungscharaktere.* I now feel that my anthropomorphisms would be easily translatable into Lewin's terms and vice versa.

10. Tolman, E. C., & Honzik, C. The effect of different degrees of hunger and of the introduction and removal of rewards on the maze performance of rats. University of California Psychological Laboratory, 1929. (In manuscript.)

11. Tolman, E. C., Robinson, E. W., & Honzik, C. The effect of different degrees of hunger upon the relative order of elimination upon long and short blinds. University of California Psychology Laboratory, 1928. (In manuscript.)

12. Washburn, M. F. Hunger and speed of running as factors in maze learning in mice. *J. Comp. Psychol.,* 1926, **6**, 151-188.

13. Williams, K. A. The reward value of a conditioned stimulus. *Univ Calif. Publ. Psychol.,* 1929, **4**, 31-55.

University of California
Berkeley, California

ON THE INHERITANCE OF MAZE BEHAVIOR

B. F. Skinner

A recent paper on the "Mode of Inheritance of Reaction Time and Degrees of Learning in Mice" by E. M. Vicari (3) purports to demonstrate the inheritance of ability to learn and the dominance or recessiveness of particular types of learning curve. The paper claims attention both because of the importance of these conclusions and because of the dimensions of the experiment itself. Miss Vicari used 900 mice. The four parent strains, referred to as Japanese waltzers, albino, dilute-brown, and abnormal-eyed, had been inbred for a considerable length of time and were apparently of great homogeneity. It was proposed to take a measure of the "reaction time" and the "degree of learning" of these strains (as given in the results with a simple maze) and to look for the presence of strain-characteristics in subsequent crosses. Three of these crosses were made, and Miss Vicari's paper deals with the resulting genetic behavior of the factors involved.

So far as the genetic manipulation is concerned, the experiment calls for no comment, unless, indeed, to point to the considerable advance which it represents over most of the earlier work on the inheritance of behavior. The supposed measures of "reaction time" and "learning," however, are open to serious objection.

Miss Vicari has employed three measures of what she terms "degree of learning." She records "(1) the total number of perfect runs made per mouse for the entire period of training; (2) the highest number of consecutive perfect runs made by each mouse for the entire training; and (3) the number of trials taken by each mouse before making the first perfect run." The data for the four parent strains are given in Table 1. The figures are means for the groups. Miss Vicari gives standard deviations of the means. These are high, averaging 89% of the mean in (1), 86% of the mean in (2), and 47% in (3).

TABLE 1

Strain	(1) Number of perfect runs	(2) Number of consecutive perfect runs	(3) Number of runs before first perfect run
Japanese waltzing	1.49	0.89	8.80
Albino	1.94	1.43	8.70
Dilute-brown	2.81	1.52	6.90
Abnormal-eyed	1.69	1.15	10.10

In interpreting the table it must be borne in mind that the type of maze was extremely simple. It involved two choices of right or left, and retracing was prevented. Purely as a matter of chance we should expect one perfect run in every four and 3.5 perfect runs for the total 14 trials. The observed number falls considerably below this in every case and averages only 1.98 for the four groups. The operation of chance, however, is not considered in Miss Vicari's treatment of her data.

We have to inquire, then, into the nature of the factors operating to deprive the animal of the successful runs which should fall to it by chance. Judging from the late occurrence of the first perfect run, practically all of the successful trials fell in the latter half of the experiment, so that the disturbance, whatever it was, diminished as the experiment progressed. [This does not threaten our contention that the perfect runs occurred by chance, since the number expected in the last seven trials (1.75) is only just exceeded by the observed number for the whole series (1.98). After the first perfect run was made, we should expect to find, in the case of the waltzers, 1.3 perfect runs from the 5.2 remaining. We get only 0.48. For the albinos we expect 1.3 in the 5.2 runs remaining, and get 0.94. For the dilute-browns we expect 1.77 in the 7.1 remaining, and get 1.81. For the abnormal-eyed we expect 0.97 out of 3.9 and get 0.69. The slight positive result for the dilute-brown stock is more than offset by the considerable shortcomings of the other observed figures.]

The only information, then, to be obtained from Miss Vicari's data is that some disturbing influence present in the experimental situation was progressively diminished during the course of the experiment. It would seem utterly unjustified to ascribe any of these conditions to "degrees of learning."

The second method of scoring consisted in measuring the time which elapsed between the animal's release and its entrance into the food box after having passed through the maze. Miss Vicari refers to this as the "reaction time." Averaging these times for each strain she is able to plot "reaction time curves" which show significant strain differences. The Japanese waltzing mice give records consistently higher than the others. Their curve falls to a minimum at the eighth run and rises slightly thereafter. The albinos, starting just under the waltzers, give a curve which descends gradually throughout the period. The other strains show no significant change in time during fourteen runs.

Although Miss Vicari's use of the phrase "reaction time" is never clear, it seems to be taken quite literally, as if the activity of the mouse in the maze were actually a reaction to a single stimulus, and as if the speed of this response were indicated in the reaction-time. She speaks of "reaction time" as the best criterion of the "maze learning response." In another instance, after finding it necessary to modify the construction of the maze because in its earlier form it encouraged her animals to "curl up in the corner and go to sleep," she says, "The time thus spent could not be counted as a fair reaction to the stimulus." She later defines this stimulus ("or urge used to induce the animal to run the maze") as "hunger." But it is obvious that hunger as a stimulus (if this has any meaning at all) could operate only after the maze had been partially learned. Actually what Miss Vicari has called "reaction time" is the lumped durations of an indefinite number of responses made to an equally indefinite number of stimuli presented by the maze situation. The nature of some of these responses will be noted later.

One thing does come out of her measurements, namely, that there are significant group differences. It would have been surprising if they had not appeared. That there are gross characteristics of behavior which show genetic constancy is, of course, common knowledge to any stock breeder. And that such gross characteristics must have influenced the "reaction times" taken in the present experiment is fully evident. But the mere inheritance of behavioral traits is so patent that there would seem to be no necessity for an experimental demonstration of the fact, unless some means of measurement were devised possessing an advantage over the method of casual inspection upon which our knowledge now rests. There is nothing in Miss Vicari's experimental method which is specifically directed toward the measurement of any given characteristic.

Frequently the investigation of the inheritance of behavior turns to the rubrics "wildness," "savagery," and "timidity." Yerkes (4) found a well-marked inheritance of these three traits, defining them in terms of the behavior of his rats in a given situation. But although he claimed a high degree of internal correlation (measuring the same animals on successive days), his criteria show no advantage over casual observation. Thus "savageness" is indicated by "(1) biting; (2) exposing or gnashing the teeth; (3) jumping at hand or forceps; (4) squeaking." Characteristics given for wildness and timidity are of an analogous sort. But the lists are not mutually exclusive and involve some fine distinctions, such as between attempts "to avoid the experimenter" (timidity) and "to hide from view" or "to escape from the hands or the forceps" (wildness). M. P. Sadownikowa-Koltzowa (2), in describing an experiment similar to Miss Vicari's, concludes that any analysis of the maze behavior of her rats must be made in terms of a multiplicity of factors. She suggests the implication of three sets of genes: "1. die Activitätsgene, 2. die Gene der Furcht und Wildheit, 3. die Gene des Ansucheninstinktes." She offers no evidence for the unity of these characteristics. Their existence is derived here, as with Yerkes, from inspection.

Nevertheless wildness, savagery, and timidity are not empty terms, and it is conceivable that they will ultimately submit to analysis and definition. Moreover, from our present knowledge of the hereditary transmission of the gross behavior which they describe, it is certain that some of their component parts show genetic constancy. The experimental demonstration of this must wait for methods of isolation and measurement. Meanwhile any attempt to show the inheritance of such a thing as ability to learn or intelligence must meet the necessity of excluding these grosser factors. With the maze this seems an insuperable difficulty.

The correlation within each group which Miss Vicari obtained (not calculated, but apparent in her curves) can be attributed to so many incidental factors showing genetic differences that we scarcely need discuss the matter further. It would appear that the reaction-times of a wild and of a tame mouse (roughly speaking) show a difference which depends upon the reactions of each animal to incidental stimuli (such as the release, the closing of doors to prevent retracing, and the like). A further difference in the records would arise from the fact that in a new situation a tame mouse runs about freely, while every movement of a wild mouse is either checked by strong inhibitions or frustrated through lack of direction. Such an influence would weaken with time and may account for the slopes of two of the present "curves." In addition to these matters of "temperament," there are others which must have affected such a measure as Miss Vicari's "reaction time." The degree of stereotropism [Crozier and Pincus (1)] shown by each strain is important, since the peculiar construction of the maze presented an array of surfaces and open spaces. Again, differences in receptors alone would account for such internal correlations against strain differences as were found. The Japanese waltzing mice, already handicapped by a difficulty in equilibration, are also deaf, and thus less responsive to auditory cues or disturbances. The abnormal-eyed are either blind or nearly so.

It is conceivable that learning might be so defined as to embrace some of these factors, but it cannot include them all without serious distortion. Miss Vicari's work, nevertheless, is offered as evidence of the inheritance of degrees of learning ability, and even of the dominance of one type of learning curve over another. It is difficult to find the warrant for these conclusions in the data presented. However careful the handling of genetic material may have been, the experiment seems irreparably damaged through the lack of an accurate and unequivocal method of measurement.

REFERENCES

1. CROZIER, W. J., & PINCUS, G. Stereotropism in rats and mice. *J. Gen. Physiol.*, 1926-27, **10**, 195-203.
2. SADOWNIKOWA-KOLTZOWA, M. P. Genetische Analyse des Temperamentes der Ratten. *Zsch. f. Induktive Abstammungs- und Vererbungslehre*, 1929, **49**, 131-145.
3. VICARI, E. M. Mode of inheritance of reaction time and degrees of learning in mice. *J. Exper. Zoöl.*, 1929, **54**, 32-88.
4. YERKES, R. N. The heredity of savageness and wildness in rats. *J. Anim. Behav.*, 1913, **3**, 286-296.

Laboratory of General Physiology,
Harvard University
Cambridge, Massachusetts

EXPERIMENTAL EVIDENCE IN SUPPORT OF PROFESSOR THURSTONE'S CRITICISM OF THE PHI-GAMMA HYPOTHESIS

MARTIN F. FRITZ

Some time ago Professor Thurstone (1) called attention to the fact that "the phi-gamma hypothesis is at best an approximation formula and that it is theoretically incorrect." He shows that if the limen is an increasing function of the stimulus then a psychometric curve must be positively skewed. Since the phi-gamma curve is symmetrical, it cannot possibly fit an actually obtained discriminal series. The discrepancy between the phi-gamma curve and the psychometric curve is rather difficult to detect in the case of fine discriminations but will be very apparent in the case of coarse discriminations.

It is the purpose of this paper to show that certain experimental data demonstrate with clearness the validity of Professor Thurstone's criticism. Evidence will be presented from two sources, (*a*) an experiment in discrimination of circles using white rats as subjects, and (*b*) an investigation of short time-intervals reported by Woodrow (2). In both investigations the discriminal functons were found to be crude.

DISCRIMINATION OF CIRCLES BY WHITE RATS

Two mature female rats were given 50 trials per day on the problem of choosing the larger of two circles. Circles (variables) of the size indicated in Table 1 were cut in heavy cardboard. Each variable was exposed

TABLE 1

Performances of a Single Rat for 11 Days, Following a 10-Day Learning Period, on the Problem of Choosing the Larger of Two Circles

The table should be read as follows: the 2-cm. variable was "judged" to be larger than the 5-cm. standard in 14 trials out of 110, or 12.7% of the time.

Stimulus magnitudes Diameter cm.	 Area sq. cm.	Number of judgments	N $v>s$ $s = 5$ cm. diam.	P $v>s$
2	3.1	110	14	12.7%
3	7.1	110	32	29.0
5	19.6	110	54	49.0
7	38.5	110	70	63.6
9	63.6	110	82	74.5

with a 5-cm. standard circle cut in the same board. By turning a board around it was possible to make a variable appear either to the right or to the left of the standard. A duplicate standard consisting of a pair of circles, each 5 cm. in diameter, was used to determine the possible existence of a constant error. The circles were exposed in front of an enclosed box containing a 60-watt lamp. The box was covered with a sheet of heavy bond paper on the side where the exposure was made in order to prevent the rats from seeing the lamp and using its position as a cue. Each rat, after running 60 cm. toward the circles, passed down through a hole in front of the circle of its choice to a runway 8 cm. below. A separate runway led from each circle to the same food box but the incorrect path was always blocked by means of a sliding door. In case an incorrect choice was made it was necessary for the rat to return, climb out of the hole, and pass down into the one before the larger circle. Only a bite of food was permitted after each trial.

All work was carried on in a dark room. The only source of light besides that in the apparatus was a small, well-shaded lamp off to one side where the records were kept. Under such conditions the rats were confronted with two distinct, well-lighted circles and with visual distractions reduced to a minimum.

Ten trials were given on each of the four variables and the duplicate standard, making a total of 50 trials per day. To keep down position habits, each group of 10 trials was so arranged that there would be five correct choices to the right and five correct choices to the left, distributed in random order. Before the experiment was begun, the rats were given 12 days of preliminary training in the apparatus in order that they might have an opportunity to become thoroughly acquainted with the general experimental setting.

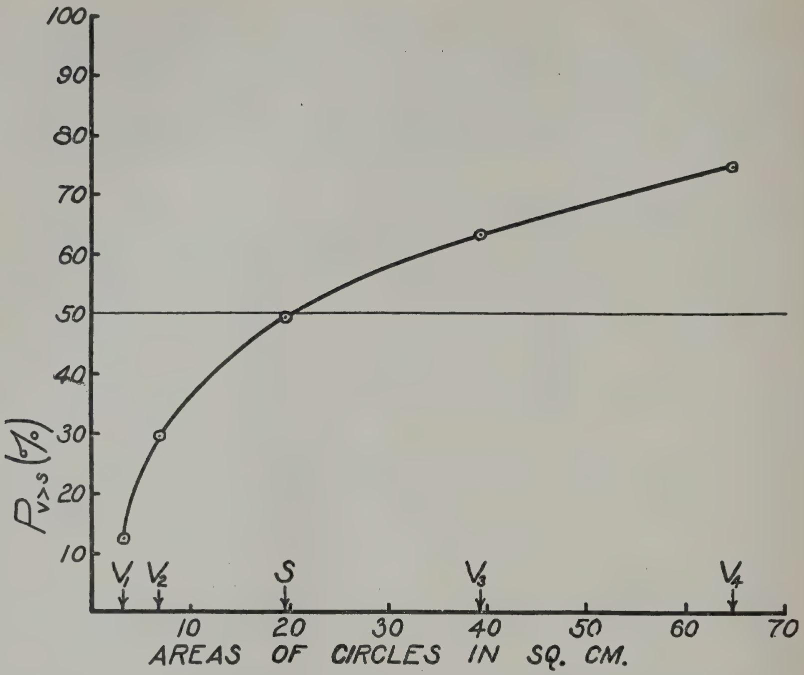


FIGURE 1

Discrimination of Circles by a Single White Rat

Graphic representation of the data of Table 1. The ordinate, $P_{v>s}$ (%), indicates the proportion "variable stimulus greater (or larger) than standard stimulus," expressed in percentage.

We should expect the early part of the experiment to be a learning period during which steady improvement would be made. Eventually, a certain level of correct discrimination or perfect discrimination would be reached, depending upon the choice of variables. Inspection of results showed that improvement occurred during the first 10 days after which there was very little increase in discriminal ability. In other words, we may consider that after about the first 10 days the rat was discriminating with nearly maximum ability.

Figure 1 shows that, in the case of a rat discriminating circles, the psychometric curve is greatly deflected to the right (positively skewed).[1] It is

---

[1] The writer is indebted to Professor Thurstone for pointing out that unless the data are the performances of a great many subjects the graphs should be made for individual cases. If the mean values of two phi-gamma curves differing greatly in precision are plotted, the resulting curve will be distorted. In order to avoid this statistical criticism, the discriminal ability of a single rat is shown in Figure 1. Since the discriminal curves of both animals are similar and positively skewed, although one showed somewhat greater accuracy than the other, it is necessary to present only the one curve.

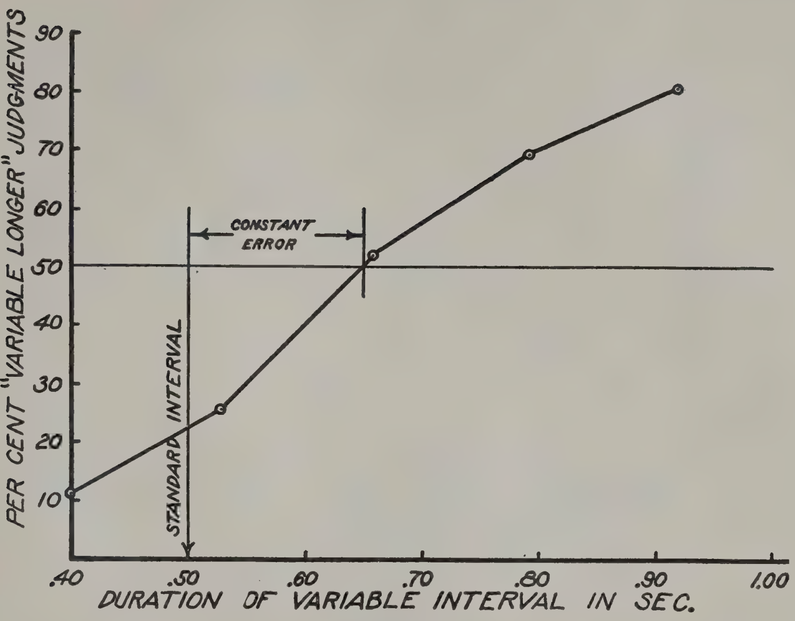


FIGURE 2

Graph Taken from Woodrow's Study of Short Time Intervals (2, p. 173)
The variable intervals were .40, .53, .66, .79, and .92 sec. in length. The percentage of "variable longer" judgments were 11.0, 25.8, 51.9, 69.6, and 81.2, respectively. The standard interval was .50 sec. There is a marked constant error. The graph is based upon 12,000 judgments secured from 30 subjects. The ordinate—percentage of "variable longer" judgments—is the same as the notation $P_{v>s}$ (%) used in Figure 1.

impossible to fit a phi-gamma curve to these experimental data. In fact, the psychometric curve obtained in this case does not even resemble the phi-gamma curve.

It is interesting to note the extremely crude discriminal ability of rats as compared to that of human beings under similar conditions. For example, discrimination of a 5-cm. and a 9-cm. circle would be no challenge to a human subject, but it seems to present a considerable degree of difficulty to the rat.

The basis of discrimination may have been either (*a*) areas of circles or (*b*) diameters of circles. The amount of light would be proportional to the areas of the circles and so reduce to the first situation, while the distances of the circles from the floor or sides of the apparatus would be proportional to the diameters and therefore reduce to the second case. It

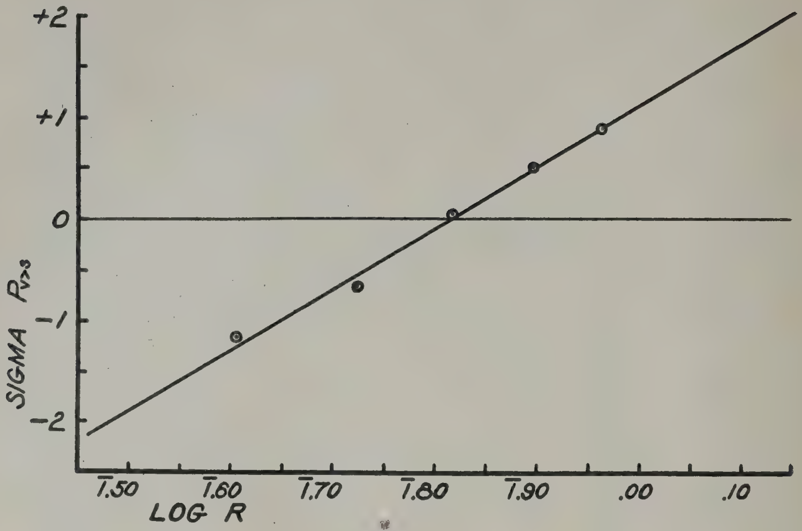


FIGURE 3

Woodrow's Data Transformed into Linear Values

makes no difference for our present purpose whether the first or the second assumption is made. All curves, whether drawn on the basis of areas or diameters, are positively skewed.

Discriminal Ability with Reference to Short Time-Intervals

Although Woodrow was interested chiefly in a study of certain constant errors which appear when short time-intervals are presented successively, his data clearly demonstrate the point with which this paper concerns itself. Discrimination is crude enough to show marked positive skewing, which again is in accordance with theoretical considerations (see Figure 2). Although the psychometric curve roughly resembles the phi-gamma curve, it is evident that a close fit between the two is impossible.

We may even go a step further and show that the amount of skewness is that which would be predicted by theory. Thurstone states that the actual fitting of the curve may be accomplished by "plotting the sigma value corresponding to the experimental proportion against the logarithm of the stimulus magnitude. The plot should be linear. A constant error will be revealed by the intercept, and the slope of the linear plot will be the precision." Figure 3 shows Woodrow's data treated in this manner. Figure 4 has been drawn by means of constants determined from the plot in Figure 3. The discrepancies between the experimentally obtained values and the theoretical values are very slight. It will be seen that this con-

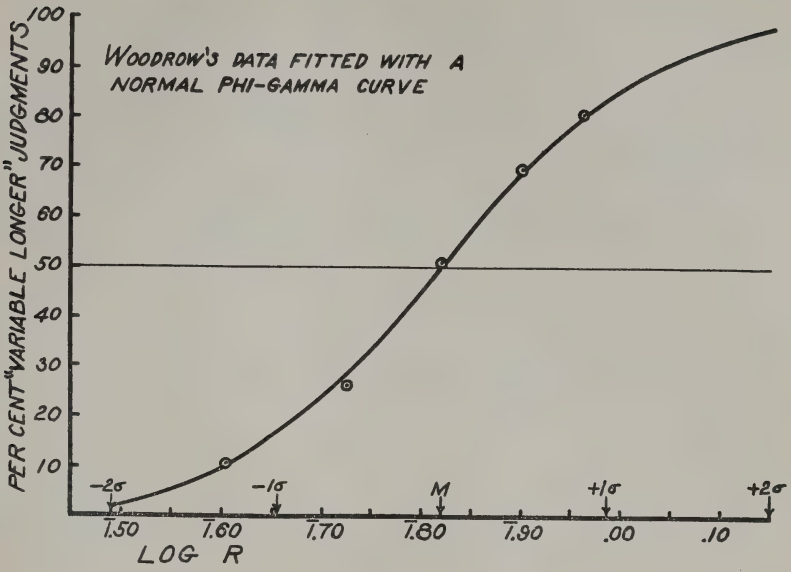


FIGURE 4

Experimentally Obtained Proportions Plotted against the Logarithms of the Stimulus Magnitudes

stitutes a very convincing argument in favor of Professor Thurstone's suggestion that, "If we want an empirical equation for the psychometric function which shall be at least not theoretically inconsistent with Weber's law, then we may propose as a substitute for the $\phi(\gamma)$ hypothesis a $\phi(\log \gamma)$ hypothesis. The integral of the normal curve would then be fitted on a psychometric function in which the proportions $p_{v>s}$ are plotted against the logarithm of the stimulus intensity....It should be regarded merely as an empirical equation for determining the limen and the precision."

Summary

1) Thurstone has demonstrated that, theoretically, a phi-gamma curve cannot be made to fit a psychometric curve. A psychometric curve is positively skewed, while a phi-gamma curve is symmetrical. The discrepancy between the two will be very apparent in the case of a coarse discrimination.

2) Experimental evidence from two sources is presented in support of this contention, (*a*) from an experiment in discrimination of circles by white rats and (*b*) from a study of short time-intervals reported by Woodrow.

3) Both investigations involve coarse discriminal abilities. The resulting psychometric curves are positively skewed and cannot be fitted by a phi-gamma curve.

4) We may correct for the amount of skewness by plotting the experimentally obtained proportions against the logarithms of the stimulus intensities. The resulting psychometric curve will be a very close approximation to the normal phi-gamma curve.

5) Theoretically and experimentally, it would seem justifiable to substitute a $\phi(\log \gamma)$ hypothesis for the $\phi\ (\gamma)$ hypothesis.

REFERENCES

1. THURSTONE, L. L. The phi-gamma hypothesis. *J. Exper. Psychol.*, 1928, **11**, 293-305.
2. WOODROW, H. Behavior with respect to short temporal stimulus-forms. *J. Exper. Psychol.*, 1928, **11**, 167-193.

Iowa State College
Ames, Iowa

AN ALTERNATIVE APPROACH TO THE MATHEMATICAL STUDY OF LEARNING CURVES

A. F. CHAISSON

The present study was undertaken, not with the object of advancing a new learning-curve equation, but rather to call attention to certain suggestions regarding the nature of learning and its mathematical statement in its quantitative aspects that were made several years ago by the biochemist, T. Brailsford Robertson, and which do not appear to have received from psychologists the attention which they undoubtedly merit.

Robertson's (5) view on learning is a by-product of his investigations into the autocatalytic nature of growth and a brief summary of his findings in this respect is therefore necessary. He has demonstrated in a rather convincing manner that the autocatalytic equation

$$log_{10}\frac{x}{A—x}=k(t—t^1) \qquad [1]$$

may be used to represent the curve obtained from growth data when weight is plotted against time. In the above equation, as applied to growth data, A is the maximum degree of growth attainable in the growth cycle and is therefore accessible to measurement; x is the amount of growth at any time, t, which is the time from the beginning of the measurements; k is a constant, specific for the process under consideration, and t^1 is the value of t when the reaction is half over, i.e., when $x=\frac{1}{2}A$. The autocatalytic equation was found to represent the growth cycle so accurately that Robertson is led to remark that such close agreement between observed and calculated values "is not frequently obtained even in experiments conducted in laboratory glassware" (6, p. 535).

Just how such a relatively simple formula which characterizes the progress of a monomolecular chemical reaction can define the time relations of such an immensely complex process as growth, and as will later be brought out, learning as well, presents something of a dilemma. We know, however, that no matter how complex a system of interrelated processes may be, the specifically slowest process sets the pace for the rest. In the case of growth, this "master reaction" is apparently autocatalytic in nature, i.e., is one of those reactions in which the products of the chemical change going on have the power of accelerating or "catalyzing" the further progress of the change until accumulation of the products of the reaction results in progressive acceleration in the opposite direction and equilibrium is finally established. The typical curve for such a reaction is S-shaped, possessing a point of inflection on either side of which the velocity relations in the two limbs of the S are different.

Turning now to the realm of central nervous phenomena, Robertson points out that their time relations are apparently governed by processes going on in the brain that are autocatalytic in nature. He demonstrates that the relationship between time and extent of movement in a simple voluntary movement such as the drawing of a straight line may be expressed by the autocatalytic equation, for when A represents the total length of the line, x the length of the line drawn in time t and t^1 the time taken to reach

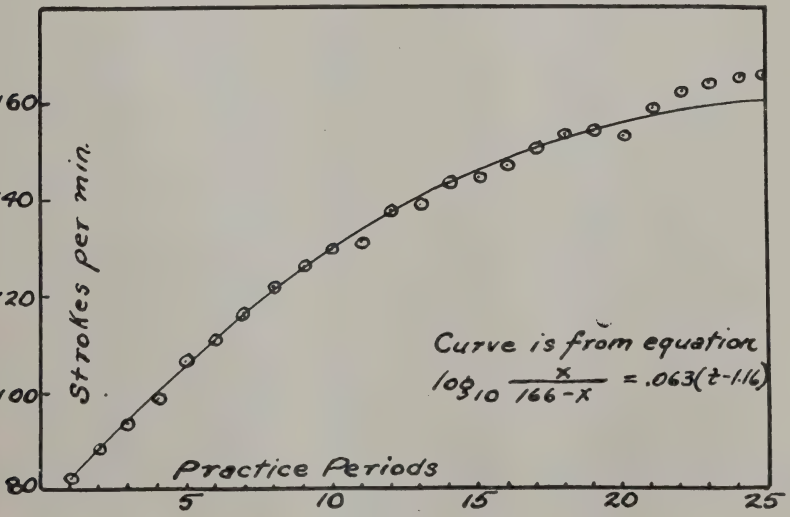


FIGURE 1

the middle of the line, k remains practically constant (6, p. 633). He also shows that the autocatalytic equation may be used to represent Ebbinghaus' data on the relation between number of nonsense syllables and learning time (6, p. 634).

The following experiment was performed by the writer in an endeavor to determine whether the autocatalytic equation could be used to represent the ordinary type of learning curve in which proficiency is plotted against time. Data for such a curve were collected by typing by the sight method a selection involving 1,668 strokes from a Gaelic translation of the *Arabian Nights*. Two practice periods were held daily, and the number of strokes per minute was calculated for each practice. The results of the experiment are shown graphically in Figure 1.

In Table 1, x is the number of strokes per minute attained at the various practices; A is the maximum number of strokes attained per minute, i.e., 166; t^1 is the value of t when $x = \frac{1}{2} A = 83$ strokes per minute. Time being reckoned in practice periods,

$$t^1=1+\frac{83-82}{88.1-82}=1.16$$

The values of $k(t-t^1)$, corresponding to values of $\frac{x}{A}$ in the curve of autocatalysis, were then determined from tables[1] with the object of finding the constant k.

TABLE 1

t	x	$\frac{x}{A}$	$k(t-t^1)$	$(t-t^1)$	k
1	82.0	.494	− .010	− 0.16	.063
2	88.1	.530	+ .052	+ 0.84	.062
3	93.9	.566	+ .115	+ 1.84	.062
4	99.0	.602	+ .180	+ 2.84	.063
5	106.5	.641	+ .252	+ 3.84	.065
6	111.0	.668	+ .304	+ 4.84	.063
7	116.6	.702	+ .372	+ 5.84	.064
8	121.8	.734	+ .441	+ 6.84	.064
9	125.1	.754	+ .508	+ 7.84	.065
10	129.3	.779	+ .547	+ 8.84	.062
11	131.1	.789	+ .573	+ 9.84	.058
12	137.9	.830	+ .689	+10.84	.064
13	139.1	.844	+ .733	+11.84	.062
14	143.9	.867	+ .814	+12.84	.063
15	144.5	.869	+ .822	+13.84	.059
16	147.2	.887	+ .895	+14.84	.060
17	150.9	.909	+1.000	+15.84	.063
18	153.8	.926	+1.097	+16.84	.065
19	154.2	.929	+1.117	+17.84	.066
20	153.0	.921	+1.067	+18.84	.056
21	151.1	.958	+1.358	+19.84	.067
22	162.3	.977	+1.628	+20.84	.078
23	164.1	.988	+1.916	+21.84	.087
24	165.7	.998	+2.698	+22.84	.118
25	166.0				

---

[1] A detailed description of the methods used in curve fitting and the set of tables used for this purpose may be found in the Appendix of (5).

It will be observed that the values of k are, as would be required by the formula of autocatalysis, almost constant. The progressive increase in proficiency in the acquisition of new Gaelic associations is therefore autocatalyzed and may be represented by the equation

$$\log_{10}\frac{x}{166-x}=0.063\,(t-1.16) \qquad [2]$$

If the observed values of $(t-t^1)$ are multiplied by the derived value of k, a series of values of $k(t-t^1)$ are obtained. From these, values of x may be calculated which are the number of strokes per minute that should be attained at the various practice periods provided that Equation [2] represents the data correctly. A comparison of observed x and theoretical x^1 values of x is made in Table 2.

It can readily be seen that the correspondence between observation and theory is very close, the only significant deviation occurring in the case of the last four readings where the rate per minute was observed to increase somewhat faster than theory demands. The fact that the autocatalytic

TABLE 2

t	$k(t-t^1)$ calculated	$\frac{x}{A}$ calculated	x^1	x	θ
1	— .010	.494	82.3	82.0	+0.3
2	+ .053	.530	88.0	88.1	—0.1
3	+ .116	.565	93.8	93.9	—0.1
4	+ .179	.602	99.9	99.0	+0.9
5	+ .244	.637	105.7	106.5	—0.8
6	+ .305	.669	111.0	111.0	0.0
7	+ .368	.700	116.2	116.6	—0.4
8	+ .431	.729	121.0	121.8	—0.8
9	+ .494	.757	125.6	125.1	+0.5
10	+ .557	.783	129.9	129.3	+0.6
11	+ .619	.806	133.8	131.1	+2.7
12	+ .683	.828	137.4	137.9	—0.5
13	+ .746	.848	140.7	139.1	+1.6
14	+ .809	.866	143.7	143.9	—0.2
15	+ .872	.882	146.4	144.5	+1.9
16	+ .935	.896	148.7	147.2	+1.5
17	+ .998	.909	150.9	150.9	0.0
18	+1.061	.920	152.7	153.8	—1.1
19	+1.124	.930	154.4	154.2	+0.2
20	+1.187	.939	155.9	153.0	+2.9
21	+1.249	.947	157.2	159.1	—1.9
22	+1.312	.954	158.4	162.3	—3.9
23	+1.375	.960	159.4	164.1	—4.7
24	+1.438	.965	160.2	165.7	—5.5
25	+1.501	.969	160.9	166.0	—5.1

equation fits this type of curve so closely renders the hypothesis highly probable that the entire curve is really S-shaped, the curve obtained experimentally being the upper limb of the S, the lower limb being represented by a basis of past associations pertaining to the specific bit of learning under consideration. This position seems to receive support from the fact that the data show a progressive decrease in the velocity of the process such as theory would demand in the case of the upper limb of the S-shaped curve of autocatalysis.

S-shaped curves of learning are actually obtainable in certain circumstances. The curves of Ebbinghaus (2, p. 65) and Meumann (4, pp. 255 ff.) showing the relation between number of nonsense syllables and number of repetitions for learning appear to be of this type. These curves, as well as one showing similar data collected by the writer, are shown in Figure 2.

Skaggs (7) has already voiced the opinion that such curves are S-shaped and contain a point of inflection on either side of which the velocity relations are different. "Just where the point of inflection occurs," he remarks,

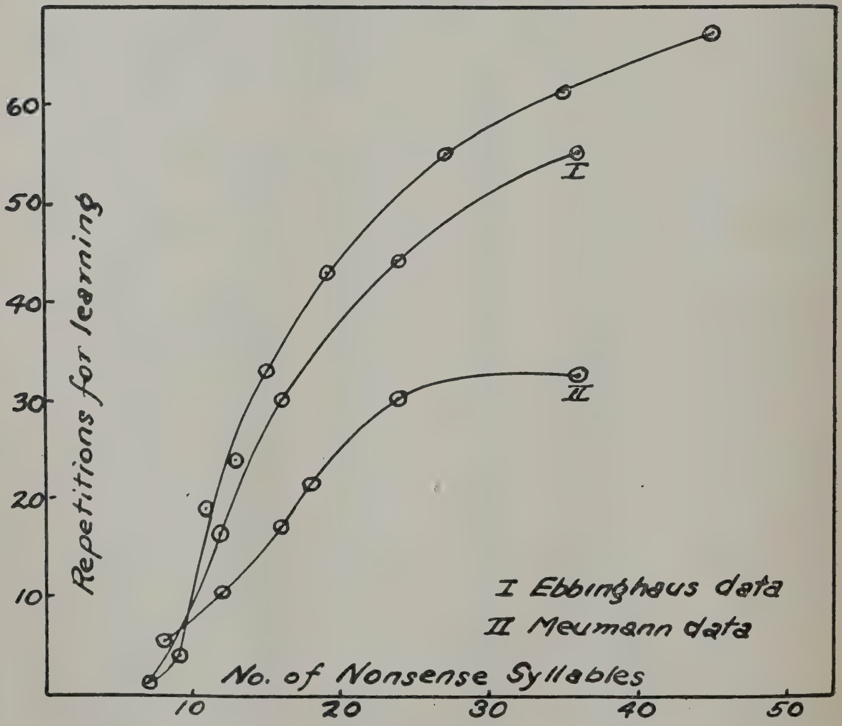


FIGURE 2

"is a question which careful experimental work alone can answer." Skaggs was evidently unaware of Robertson's work on Ebbinghaus' curve, which undoubtedly furnishes evidence in support of his position, for Robertson shows that Ebbinghaus' data may be expressed by the autocatalytic equation

$$\log_{10}\frac{x}{43.6-x}=0.001486\,(t-0.526)\qquad[3]$$

and that the only significant difference between theory and observation occurs in the case of the number of syllables which may be learned in a single repetition (6, p. 634). In Table 3, Ebbinghaus' observations are shown, together with Robertson's calculations from Equation [3], t being the total learning time for the series of syllables, r the number of repetitions, v the number of repetitions per syllable, x the observed number of nonsense syllables, and x^1 the number of syllables calculated by Robertson from Equation [3].

TABLE 3

t	r	v	x	x^1
2.8	1.0	0.14	7	10.1
79.7	16.6	1.39	12	12.2
192.0	30.0	1.88	16	15.8
422.0	44.0	1.83	24	24.2
792.0	55.0	1.53	36	35.4

Robertson's value of A in Equation [3] would mean that t^1 would fall at 21.8 syllables. The values of v in the above table seem to furnish evidence that this point really does represent a point of inflection in the curve for the number of repetitions per syllable increases up to this point and decreases beyond it.

These findings have been checked against data on the memorization of nonsense syllables collected by the writer. The method used was that of Ebbinghaus and involved the reading and rereading of series of nonsense syllables at a definite rate until they could be repeated without error. The time spent in the perusal of each syllable was .50 second and the total length of the reading periods was made the same by the inclusion of a sufficient amount of other material. Results are shown in Table 4 and k was determined in the same way as in the case of the Gaelic curve. A was taken as 45, and t^1 was 554.6 seconds.

The calculated values of x (i.e., x^1) have been obtained from the equation

$$\log_{10}\frac{x}{45-x}=0.001059\,(t-554.6)\qquad[4]$$

and their differences from the observed values are shown in the column

TABLE 4

t	r	v	x	$\frac{x}{A}$	$k(t—t^1)$	$(t—t^1)$	k	x^1	θ
3.5	1	0.14	7	.155	—.737	—551.1	.00132	9.4	+2.4
18.0	4	0.44	9	.200	—.602	—536.6	.00112	9.6	+0.6
104.5	19	1.73	11	.244	—.491	—450.1	.00108	11.3	+0.3
156.0	24	1.84	13	.288	—.393	—398.6	.00098	12.4	—0.6
247.5	33	2.20	15	.333	—.302	—307.1	.00098	14.3	—0.7
408.5	43	2.26	19	.422	—.137	—146.1	.00093	18.6	—0.4
742.5	55	2.03	27	.600	+.176	+187.9	.00094	27.5	+0.5
1067.5	61	1.74	35	.777	+.542	+512.9	.00106	34.9	—0.1
1507.5	67	1.48	45	1.000	+.000	+952.9		40.9	—4.1
				Mean k=.001059					

on the extreme right. Equation [4] represents the findings quite accurately except in the case of the two extremes. If the maximum *A* were determined more accurately it is possible that the fit would be even closer, 45 syllables being but an approximation of the maximum of what appears to be a first cycle of the curve (7, p. 152). A point of inflection seems to exist in the curve, since the number of repetitions per syllable (v) is seen to increase on one side of our value of t^1 and to decrease on the other. These findings render the conclusion highly probable that the autocatalytic equation may be used to express the relation between length of task and learning time, at least as far as the learning of nonsense material is concerned.

CONCLUSIONS

1) Robertson's view that the time relations of central nervous phenomena are governed by processes taking place in the brain that are autocatalytic in nature receives support from our experiments.

2) The autocatalytic equation has been found to accurately express a typical learning curve in which proficiency is plotted against time.

3) The same equation was found to represent the relation between length of task and learning time in the case of nonsense material. Skaggs' belief that the curves of Ebbinghaus and Meumann showing the relation between number of nonsense syllables and number of repetitions for learning are S-shaped and possess a point of inflection is supported by our findings. His "suspicion" that when the number of syllables in the series becomes very long, a second S-shaped cycle will be superimposed upon the first appears highly probable.[2]

---

[2] Cf. the following statement by Robertson: ". . . . a variety of measurements which have been made on the rate of learning by telegraph operators, typists and so forth, all yield 'curves of learning' which very strikingly resemble the curve which represents the progress of an autocatalysed chemical reaction, and in some cases, it appears, two or more of such curves may be superimposed to yield 'cycles of learning' just as we have cycles of growth in the growing organism" (6, p. 635).

4) The fact that the autocatalytic equation represents the curve for the formation of Gaelic associations so accurately suggests the hypothesis that curves of this type are the upper limbs of S-shaped curves of autocatalytic nature.

5) Since the time relations of both growth and learning appear to be governed by the same type of reaction, this common element in the two is of considerable interest in view of the work of Kappers (3) and Bok (1).

6) While it is by no means claimed that the autocatalytic equation is the only equation that will adequately express learning data, it does seem to account more readily for observed facts and to be of wider application than any other equation thus far proposed and of which the writer is aware.

REFERENCES

1. Bok, S. T. Stimulogeneous fibrillation as a cause of the structure of the nervous system. *Psychiat. en Neur. Bladen,* 1915, **19**, 393-408.
2. Ebbinghaus, H. Ueber das Gedächtnis: Untersuchungen zur experimentellen Psychologie. Leipzig: Duncker, 1885. Pp. 169.
3. Kappers, C. U. A On structural laws in the nervous system: the principles of neurobiotaxis. *Brain,* 1921, **44**, 126-149.
4. Meumann, E. The psychology of learning. An experimental investigation of the economy and technique of memory. (Trans. from the 3rd German ed. by J. W. Baird.) New York: Appleton, 1913. Pp. xix+393.
5. Robertson, T. B. The chemical basis of growth and senescence. Philadelphia: Lippincott, 1923. Pp. xiii+389.
6. ———. Principles of biochemistry for students of medicine, agriculture and related sciences. (2nd rev. ed.) New York: Lea & Febiger, 1924. Pp. xii+796.
7. Skaggs, E. B. The relation of length of material and number of repetitions needed to learn. *J. Gen. Psychol.,* 1929, **2**, 150-152.

St. Francis Xavier University
Antigonish, Nova Scotia

A STUDY OF LEARNING CURVES: II. THE RELATIONSHIP BETWEEN THE HYPERBOLA AND THE ARC COTANGENT FUNCTION[1]

Grau

Willard L. Valentine

Ettlinger (1, 2) has shown that the hyperbola

$$t=t_l\frac{x+a}{x} \qquad [1]$$

as used by Thurstone (4) to fit learning data, is a first order approximation to the growth curve,

---

[1] This study was made possible through a grant from the Graduate School of the Ohio State University.

$$t = \frac{t_l}{1 - e^{-\frac{x}{a}}} \qquad [2]$$

The difference between the two curves is expressed by

$$\frac{2\left(1 + \frac{x}{a e^{\frac{x_1}{a}}}\right)}{e^{\frac{x}{a}}}$$

which, when evaluated, is smaller than the experimental error in collecting the data and fitting the curves to them. In both [1] and [2] x is some measure of the amount of practice and t is time required to perform a unit act.

We propose to show that Meyer's (3) arc cotangent function also approximates the hyperbola whenever x is beyond some small positive number. This approximation has been shown to hold in specific cases [Valentine (5)].

Thurstone's equation may be written

$$t = t_l \frac{(x+p)+a}{x+p} \qquad [3]$$

and Meyer's [Valentine (6)] as

$$t = t_l + k \cot^{-1} s(x+p) \qquad [4]$$

in both [3] and [4] p is a measure of equivalent previous practice and t_l is the limit of both as x approaches infinity. By transposing and writing t' for $t - t_l$ and x' for $x+p$

$$t' = t - t_l = \frac{t_l a}{x'} \qquad [5]$$

$$t' = t - t_l = k \cot^{-1} s x' \qquad [6]$$

To find the values of the parameters of [6] so that the curve is tangent to [5] at the point $[(t_l a)^{½}, (t_l a)^{½}]$, let $(t_l a)^{½} = b$, and, since the curve must pass through (b,b),

$$b = k \cot^{-1} s b \qquad [7]$$

or

$$k = \frac{b}{\cot^{-1} s b} \qquad [8]$$

The first derivatives of [5] and [6] are

$$\frac{dt'}{dx'} = -\frac{b^2}{x^2}$$

and

$$\frac{dt'}{dx'} = \frac{ks}{1+s^2x'^2}$$

$$\therefore \quad 1 = \frac{ks}{1—b^2s^2}$$

or

$$b^2s^2—ks+1=0$$

and

$$s = \frac{k\pm\sqrt{k^2—4b^2}}{2b^2}$$

Substituting the value of *k* from [8] and transposing

$$2bs = \frac{1}{cot^{-1}sb} \pm \sqrt{\frac{1}{cot^{-1}sb} - 4} \tag{9}$$

Solving for sb, let $v = cot^{-1}sb$, then after substituting in [9] and simplifying, we have

$$cot^2v+1 = csc^2v = \frac{1}{v}\, cot\ v$$

$$1 = \frac{1}{v}\, cos\ v\ sin\ v$$

Multiplying by sin^2v, we have

$$2v = sin^2v,$$

but this is true only when $v=0$.

But from [7] $k = \frac{b}{v}$; so $s = \frac{cot\ v}{b}$

Substituting in [6]

$$t' = \frac{b}{v}\, cot^{-1}\left[\frac{cot\ v}{b}\, x'\right] \tag{10}$$

As $v\rightarrow 0$, [10] approximates [5] which may be verified by considering that, for small angles, the cotangent of the angle is approximately equal to the reciprocal of the same number, or

$$\underset{v\rightarrow 0\ app.}{cot\ v} = \frac{1}{v}$$

$\therefore$ from [6]

$$t' \underset{app.}{=} \frac{b}{v}\, cot^{-1}\left[\frac{x'}{bv}\right]$$

or

$$t' \underset{v\to 0}{=} \frac{b}{v}\cdot\frac{bv}{x'} = \frac{b^2}{x'} = \frac{t_l a}{x'} \qquad \text{from [5] and [7]}$$

which is the equation of the hyperbola in t'.

Therefore, we may conclude that for all positive values of x beyond some point slightly above zero, the arc cotangent function may be made to approximate the hyperbola closer than either curve will approximate the learning scores.

REFERENCES

1. ETTLINGER, H. J. A curve of growth designed to represent the learning process. *J. Exper. Psychol.*, 1926, **4**, 409-414.
2. ————. The kinetics of learning. *Amer. Math. Mo.*, 1926, **33**, 506-510.
3. MEYER, M. F., & EPPRIGHT, F. O. The equation of the learning function. *Amer. J. Psychol.*, 1923, **34**, 203-222.
4. THURSTONE, L. L. The learning curve equation. *Psychol. Monog.*, 1919, **26**, No. 114. Pp. 51.
5. VALENTINE, W. L. A graphical method for fitting curves to learning data. *J. Gen. Psychol.*, 1930, **3**, 337-340.
6. ————. A study of learning curves. I. The application of Meyer's arc cotangent function and Thurstone's hyperbola to the maze performance of white rats. *J. Comp. Psychol.* In press.

The Ohio State University
Columbus, Ohio

TIME ESTIMATION IN THE HYPNOTIC TRANCE[1]

JOHN M. STALNAKER AND MARION W. RICHARDSON

Several apparently trustworthy observers of hypnotic phenomena have noted an increased ability in post-hypnotic estimation of time. Inasmuch as practically all post-hypnotic phenomena may be produced with greater strength during the trance state, the following experiment was undertaken to test the accuracy of the estimation of short intervals of time in the trance state.

The subjects, nine in number, were all what is commonly called well-trained trance subjects. Each of them had been used for demonstrational work at least four or five times before the experimental work was undertaken, and had demonstrated such phenomena as negative hallucinations, automatic writing, analgesia, catalepsia, and amnesia. In the subsequent waking state all of them showed complete verbal-recall amnesia for trance events.

---

[1] The experiment here described was devised and carried out by J. M. Stalnaker. The statistical analysis of the results was done by M. W. Richardson.

The original method of trance induction was optical fixation, but, as the subjects became more accustomed to the state, they were put into the trance by direct verbal suggestion, aided by post-hypnotic suggestions from the previous trance state.

During each experimental period the subject was asked to estimate 12 time intervals, 6 in the waking, and 6 in the hypnotic state. Each subject was used for 10 experimental periods. Every effort was made to keep the conditions within the room uniform. Noises from the outside were never great, although they were somewhat variable.

The intervals of time used for estimation were one, two, and three minutes. One day the subject began estimating in the waking state, while the next day he began in the trance state. He estimated these intervals in the order given below. The procedure was the same in both waking and trance states, except that different chairs were used. Directions to the subject differed only in one respect: in the trance the subject was told that his estimations would be more accurate. He was directed in both states not to use any aids to estimation such as counting, noting the pulse, or other similar device. This was found to be necessary after one or two subjects, tried out in advance of the experiment, were found to have spontaneously used such devices with astonishing success. The subject was directed to let his mind be as free as possible, was informed that the starting signal would be the "three" of a slow uniform one-two-three count aloud by the experimenter. The subject was instructed to indicate his estimation by saying "time" at the expiration of what he believed to be the correct interval. The directions were readily understood by all subjects.

Two decimal stop watches were used for the timing. Their variation was always less than 1/100 of a minute.

Two different series of time interval arrangements were alternated. The first series upon which half of the subjects started was:

1) Waking state. Estimation of 1, 3, and 2 minute intervals.
2) Trance state. Estimation of 2, 1, 3 minute intervals.
3) Trance state. Estimation of 2, 3, 1 minute intervals.
4) Waking state. Estimation of 3, 1, 2 minute intervals.

The subject was put into the trance state as soon as he had finished the waking estimations (1). The trance estimations (2) and (3) were taken in order without any lapse of time between. Following the completion of the trance estimations, the subject was awakened. He was given a post-hypnotic suggestion to awaken at once, and *apparently* the subject did just that.

On alternate days the procedure was as follows:

1) Trance state. Estimation of 3, 1, 2 minute intervals.
2) Waking state. Estimation of 1, 3, 2 minute intervals.
3) Waking state. Estimation of 2, 1, 3 minute intervals.
4) Trance state. Estimation of 1, 2, 3 minute intervals.

The subjects were used in this work one at a time. Only the subject and the experimenter were present. In general the subjects appeared each day at the same hour. They did not object to the experiment in the least, and, in general, they believed that they estimated very accurately.

RESULTS

In Table 1 the results for all subjects are lumped together, since analysis of the individual records shows that, with the exception of G, described below, no noteworthy differences are found. It is clear from Table 1 that no significant differences in the average error in estimating these time intervals are found.

TABLE 1

SUMMARY OF THE ERRORS OF ESTIMATION OF ALL SUBJECTS IN BOTH TRANCE AND WAKING STATE, TOGETHER WITH DIFFERENCES, STANDARD ERRORS, ETC.

	Trance			Waking			Difference			
Time interval	Mean error in estimation	*S.D.*	σ_m	Mean error in estimation	*S.D.*	σ_m	*T-W*	*W-T*	$\sigma_{diff.}$	$\frac{Diff.}{\sigma_{diff.}}$
1 min.	.402	.332	.027	.397	.353	.028	.005		.039	0.13
2 min.	.608	.428	.035	.682	.504	.041		.074	.053	1.2
3 min.	.902	.728	.059	1.000	.718	.057		.098	.082	1.2

This table is interpreted thus: for the one-minute interval, the mean error in estimation, disregarding sign, was .402 minutes, for the trance state. The number of estimates (not given in the table) is 154, since two records are incomplete. The standard deviation of these 154 errors is .332 minutes. The reliability of the mean error is indicated by the sigma of the mean, .027 minutes. The next three numbers, reading horizontally, are similar measures for the 154 waking estimations of one-minute intervals. The slight difference in the average error in estimation is .005, slightly greater in the trance state. The $\sigma_{diff.}$ of .039 makes this difference quite unreliable, since $\frac{Diff.}{\sigma_{diff.}}$ is only 0.13.

Only one subject, G, shows results that might be suggestive of specific effects of the trance state on the estimation of time. In Tables 2 and 3, his estimates are considered separately. Table 2 shows that the results of G agree with the pooled data of all the subjects in that there is no significant difference in the average error of estimation. Table 3, however, shows that he has a distinct tendency to make the estimates too long in the trance state and too short in the waking state. Table 4 shows that this is not true for the pooled data; moreover analysis of individual records did not reveal this tendency in any other subject. It must be stated that G was an excellent subject, as judged by the customary criteria of the depth of trance, but not better than most of the other subjects. It may be interesting to note that he had a high degree of subjective certainty about the accuracy of his estimations. He was quite sure that all of his estimations were practically perfect.

TABLE 2

SUMMARY OF THE ERRORS OF ESTIMATION FOR SUBJECT G, IN BOTH WAKING AND TRANCE STATES

	Trance		Waking		Difference			
Time interval	Mean error in estimation	σ_m	Mean error in estimation	σ_m	*T-W*	*W-T*	$\sigma_{diff.}$	$\frac{Diff.}{\sigma_{diff.}}$
1 min.	.217	.116	.123	.071	.093		0.135	.69
2 min.	.398	.199	.641	.301		.243	0.361	.67
3 min.	.558	.277	1.051	.503		.493	0.574	.86

Table 2 is to be interpreted in the same way as Table 1. The number of estimates by Subject G is 20 in each of the two states for each of the time intervals indicated.

TABLE 3

THE AVERAGE TIME TAKEN FOR THE ESTIMATION OF THE THREE INTERVALS BY SUBJECT G

	Trance		Waking				
Time interval	Mean estimate	σ_m	Mean estimate	σ_m	Difference *T-W*	$\sigma_{diff.}$	$\frac{Diff.}{\sigma_{diff.}}$
1 min.	1.217	.034	.953	.034	0.264	.049	5.4
2 min.	2.398	.052	1.410	.085	0.988	.099	9.9
3 min.	3.468	.098	2.007	.140	1.461	.171	8.5

Table 3 is interpreted thus: the one subject G made an average estimate of 1.217 minutes for the one-minute interval during the trance state; the sigma of this average is .034 minutes. During the waking state his 20 estimates of the one-minute interval averaged 0.953 minutes, with a sigma of 0.034 minutes. The average trance estimate was .0264 minutes greater than the average waking estimate; this difference is wholly reliable as indicated by the $\frac{Diff.}{\sigma_{diff.}}$ of 5.4, etc.

Further study of the experimental data showed that no noteworthy individual peculiarities existed, except those of G, and that the experimental conditions seemed to have controlled practice and accidental factors very well.

TABLE 4

THE AVERAGE TIME TAKEN FOR THE ESTIMATION OF THE THREE INTERVALS BY ALL SUBJECTS

	Trance			Waking					
Time interval	Mean estimate	*S.D.*	σ_m	Mean estimate	*S.D.*	σ_m	Difference *T-W*	$\sigma_{diff.}$	*Diff.* / $\sigma_{diff.}$
1 min.	1.201	.490	.040	1.132	.515	.042	.069	.057	1.2
2 min.	2.082	.739	.060	1.988	.848	.068	.094	.091	1.0
3 min.	3.043	.585	.047	2.830	1.215	.098	.213	.106	2.0

The σ_m, of course, of Tables 3 and 4 are based on a number of observations too limited for adequate statistical investigation.

CONCLUSIONS

1) Within the limits of the experimental conditions and the procedure outlined above, there is no difference in the average error of estimation of time intervals between the trance and waking states.

2) There is no significant tendency to make the estimate too large or too small that does not exist in both trance and waking states.

3) One subject differed from the group in showing the tendency to make the estimations during the trance too large, and the estimations in the waking state too small.

Purdue University
Lafayette, Indiana

A QUANTITATIVE METHOD OF TESTING THE DIMINUTION OF THE EFFECT OF POST-HYPNOTIC SUGGESTION WITH THE PASSAGE OF TIME[1]

JOHN M. STALNAKER AND E. W. CUMMINS

One of the most significant and characteristic features of hypnosis is found in the post-hypnotic phenomena. The subject, while in the trance state, is told that he will perform a specific act at a given signal after awakening. The subject after he awakens normally has verbal-recall amnesia for the trance events; that is, he says he cannot remember anything that happened while he was in the trance, but at the designated signal he carries out the act just as he was told he would. He either rationalizes, when asked why he behaved in this way, or declares he did not realize that he executed the act; i.e., he has amnesia for the specific act.

---

[1] All hypnotic work in this experiment was done by John M. Stalnaker. Indebtedness is hereby expressed to Professor Clark L. Hull and to Professor W. S. Taylor for creating an interest in the problems of hypnosis. Neither of these men, however, is in any way responsible for the procedure herein described.

A quantitative study of the nature of the strength and duration of post-hypnotic suggestion is a vital problem both to a better scientific understanding of hypnosis and to a more complete use of its therapeutic powers.

It is the purpose of this paper (*a*) to describe a technique by which a quantitative study of the diminution of post-hypnotic suggestion may be made, and (*b*) to give the actual data obtained from the use of this procedure.

In selecting a method to test the diminution of the effect of post-hynotic suggestion during a lapse of time, several important considerations are necessary. The method must permit that a quantitative measure of the diminution be taken. The suggestion given must be of such a nature that the subject will have no knowledge of its operation. That is, the suggestion should not demand the execution of an act so foreign to the habits of the subject that he will, through a process of logical reasoning, attribute it to the effects of the experiment. This precaution should be observed, since a knowledge, real or supposed, of the suggestion given may influence its operation. The suggestion should not require the execution of an act which may readily tend to become habitual. Of course, should a strong habit be set up, there would no longer be a measure of diminution of the effect of the suggestion, since the execution of the act would have passed out of the jurisdiction of the suggestion. No laboratory equipment is required in this experiment.

The basic idea of this procedure is to use hypnosis to direct the subjects to hand in each day a report blank giving several items of information, among others, the exact hour at which they awoke. This notion is given them both in the waking and the trance state several times, and they are trained to fill out regularly the report blanks and to turn them in. Then, while the subject is in a deep trance state, the suggestion is given very clearly and strongly that he will awaken every third day, starting with the next day, at exactly 4 o'clock; note the time; and return to sleep. The remainder of the experiment consists in an examination of the record blanks, the subjects not being used as trance subjects again.

Arrangements are made that the subjects will appear for the experiment at the same time each day for four successive days. The first three days are devoted to preliminary instructions, while on the fourth day only is the test suggestion given.

The first subject—the subjects are treated individually—is placed in a trance. The subjects used should all be what is commonly termed good trance subjects who have been put in a trance state by the operator before this meeting. After the operator is convinced that the subject is in his usual depth of trance state, directions are *read* to the subject to the effect that he is to follow the directions very precisely, that the task will require no appreciable amount of time or effort on his part, and that he will enjoy the work. He is then informed that he will keep an accurate account of some of his daily habits and that to facilitate this a blank has been prepared with spaces which he is to fill out each day. Immediately after supper each day, starting the very next day, he will fill out one of these blanks as accurately as is possible. This operation will take only a few seconds. He will never forget to do this. In fact, he will feel uncomfortable and ill-at-ease following supper until the blank is filled. The following day the report will be turned in at a given place. The matter is to be kept entirely secret.

After these instructions have been read, the usual suggestions preliminary to awakening a subject are given—i.e., he will feel fine when he awakens, he will be a better subject next time, he will enjoy cooperating, he will feel relaxed, refreshed, etc.

After the subject has come out of the trance, of course, he does not recall the directions given him in the trance. These directions are therefore repeated to him in the waking state. A supply of blanks is given to him. The information requested on the blank is as follows:

Name .... Date ....

1. How did you spend the major part of last evening? .... ________
2. Time of retiring last night .... ________
3. Time you first awoke this morning .... ________
4. The time you arose .... ________
5. How did you feel when you got up? *a.* v. good; *b.* good; *c.* average; *d.* badly; *e.* v. badly ... ________
6. Did you sleep well? .... ________
7. Hours of meals .....Breakfast .... ________
   Lunch .... ________
   Supper .... ________

Please fill out this report accurately immediately after supper each day, and turn it in the following day. Do not show it to anyone. This is confidential.

On the next day and the day following, the subject returns at the same hour and is again put into a trance state. A modified set of directions is *read* to him on each of these days. He is asked, while in the trance state, if he is finding any difficulty in complying with the directions. The same points stressed in the first day's suggestions are again stressed. In the waking state he is again told to continue filling out and handing in the blanks.

Upon the fourth day the subject is again placed in the trance state, this time particular effort being used to insure that he is in a very deep state. The following suggestion is then read: "Each third morning hereafter, beginning with tomorrow morning, Friday morning, you will awaken *exactly* at 4 o'clock, note the time, which you will remember when you fill out your report, and then fall soundly asleep immediately afterwards. Saturday and Sunday mornings you will awake at your accustomed hour, but on Monday, the third morning after Friday, you will awaken again exactly at 4 o'clock. Then Tuesday and Wednesday will be normal, with Thursday the third day, the day upon which you awaken at exactly 4 o'clock." This suggestion is read again very carefully and slowly to the subject. The following additional suggestions are then given: "After you fall asleep again on these third mornings which you *will do* very quickly, you will sleep soundly and awake at your accustomed hour feeling fine. The sleep you get between 4 o'clock and the next time you awaken will be a very sound and deep sleep from which you will greatly profit. Thus, when you awake later, you will feel fine.

"Now let me make this perfectly clear. From tomorrow morning on without any further suggestion you will wake up each third morning at exactly 4 o'clock, make a mental note of the time, which you will remember when you fill out your daily report, and then return to sleep. You will continue, without any further suggestion, to hand in your reports just as you have been. You will not be used as a trance subject for sometime, but the suggestions which I have given you will carry out until you are informed in hypnosis that they will no longer be effective. Do you understand this?"

The subject is expected to reply in the affirmative, but if he does not understand the directions, they are made perfectly clear to him. The subject is then awakened in the usual fashion; he is told in the waking state to continue handing in his reports just as he has been, but that until he is notified he will not need to report for any further experimental work.

Control subjects should be used who do not go into a trance state. The directions should be given to them in precisely the same fashion as they are given to the experimental subjects.

It will be seen that this procedure conforms to the considerations set forth above. A quantitative measure of diminution of effect of the suggestion can be made by recording the waking times of the days upon which the subjects were instructed to awaken at 4 o'clock, and since it was suggested that they would forget the fact that they awoke at 4 o'clock except at the time they filled out their reports, they would have no knowledge of the operation of the suggestion. This action is not likely to become habitual for two reasons; first, because of its disagreeable nature; and, secondly, because the third-day arrangement does not fit in with any other time measurement, such as the day of the week, the alternate-day arrangement of classes, etc.

In order to test out this procedure to see if it was actually workable, six subjects who had evidenced a special interest in hypnotism were selected. Three of these were excellent hypnotic subjects and had been used for demonstrational work several times. They exhibited in perfectly orthodox fashion all of the usual symptoms of a deep trance. The three control subjects were students who had tried to go into a trance state several times but in no case had they shown any effects of such a state.

The procedure as described above was begun at 4 P.M. Thursday, February 12, 1929. At this time the first subject received the first suggestion. The remaining subjects reported at 20-minute intervals. Each subject was treated exactly in the same manner. The control subjects were asked to relax as completely as possible, to close their eyes, to make themselves as susceptible to suggestions as possible, and to listen carefully to the suggestions. In the case of the hypnotic subjects they were placed in a trance by one of the writers—J. M. Stalnaker—while the directions were read to them by the other writer—E. W. Cummins. The ordinary precautions in effecting a transfer of rapport were followed. After the suggestions had been read, a transfer of rapport was again effected, and they were awakened from the trance state by the same operator who had put them in.

It is unfortunate that the regular closing of school necessitated bringing this experiment to a premature conclusion. On May 28, 1929, 105 days after the initial suggestion had been given, the suggestion was removed while the subject was in the trance state. While the hypnotic subjects were in this final trance state, and before the suggestion had been removed, each subject was asked the following questions:

1) Have you experienced any mental or physical discomfort during this experiment which you thought was due in some manner to the suggestions?
2) Have you formulated any logical opinions as to the nature or purpose of this experiment?
3) Would you object to continuing this experiment for several months more?

These questions were also asked of each subject in the waking state following the removal of the suggestion in the final trance state. The answers in both trance and waking state agreed. There was experienced no mental or physical discomfort. No logical opinions were formulated. None of the subjects objected in the least to continuing the experiment for several months more. The control subjects, of course, believed that they knew what the purpose of the experiment was, inasmuch as they could clearly recall all of the directions which had been given them.

A check was made to ascertain whether or not the subjects were actually awakening at the hour they stated. It was found that they were. This was indicated by remarks the subjects made to their associates, and, in one case, by observation by a roommate. As far as the writers are aware, however, no cases of direct lying have been found in deep hypnosis. Hence, when the subjects in the trance state say, as these subjects did, that they

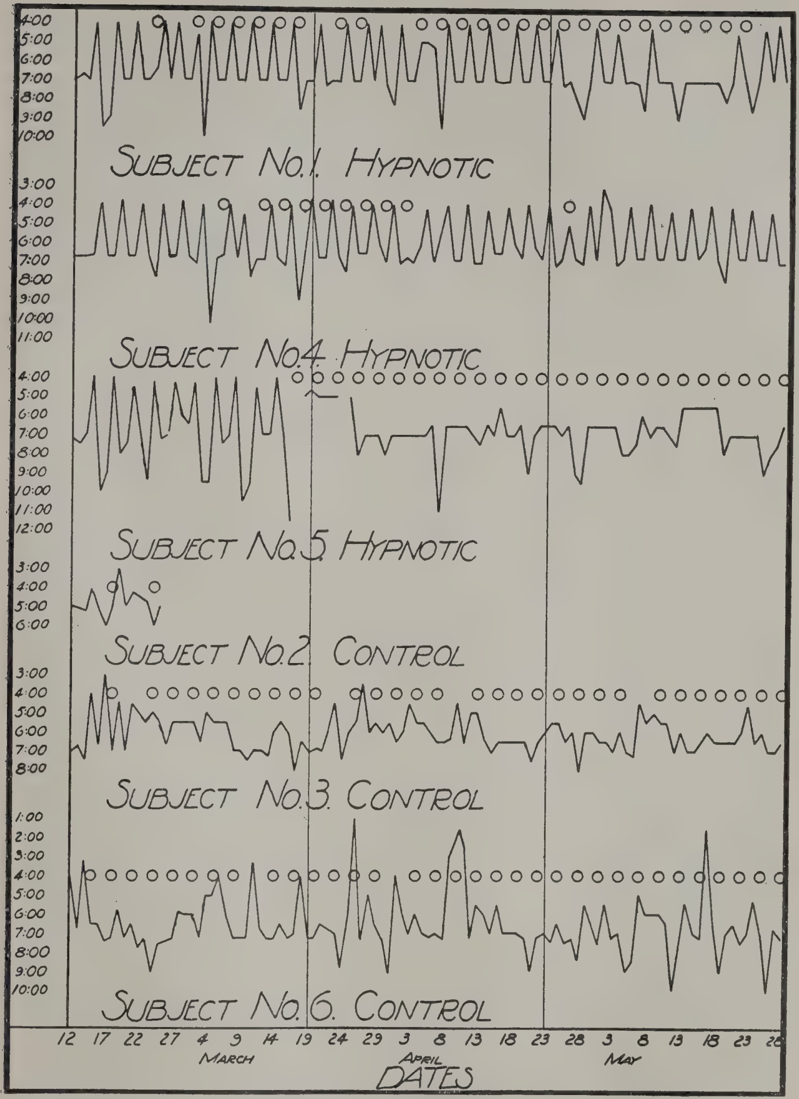


FIGURE 1

THE WAKING TIMES OF THE HYPNOTIC AND CONTROL SUBJECTS

The suggestion was given on March 14 that the subjects should awaken every third morning at exactly 4 A.M. The circles indicate the days upon which the subjects did not awaken as instructed. It may be noted, in several instances, that the hypnotic subjects carry out the suggestion, even though it is a day late.

awoke at a certain hour, that is probably true. Even should the behavior being measured, be the act of writing down, every third day, 4 A.M., the procedure herein described would still be valid.

The accompanying graph gives a detailed picture of the results. Little interpretation is necessary. The first hypnotic subject (No. 1) shows some irregularity, but the potency of the suggestion has by no means lost its power even at the end. The second hypnotic subject (No. 4) started one day late with the suggestion, but there appears to be no tendency for the suggestion to lose its effect. In the case of the third hypnotic subject (No. 5), the suggestion faded away apparently completely at the end of 30 days. This subject was initiated into a fraternity at this time. He was kept awake all night on March 17 and again on March 24. The initiation process seemed to cause the suggestion to fade out. There seems to be no reappearance of the suggestion after this time.

The graphs of the control subjects show in no case any tendency to awake according to the suggestion. Subject No. 2 got up early regularly in order to fire furnaces. He was taken ill and dropped from the experiment, as is indicated on the graph.

No general conclusions are justified because of the small number of cases used here and because of the premature termination of the experiment. It is safe to say, however, that the technique described here is a workable one. Inasmuch as no two subjects show exactly similar results, more subjects will have to be used. The suggestion appears to operate with regularity and strength. No doubt two methods of measuring the fading-out effect of the suggestion may be necessary in this experiment: first, by the subjects awakening later than 4 A.M. on the specified mornings, but earlier than usual, and, secondly, by the subjects failing to awaken on the proper day, but awakening the day before or the day after at the correct hour. This "skipping" of days may become pronounced with a lapse of time. The behavior of subject No. 5 indicates that unusual happenings of an emotional nature, such as a fraternity initiation, may destroy the suggestion.

Purdue University
Lafayette, Indiana

TIME FACTORS IN CONDITIONING FINGER-WITHDRAWAL*

HELEN MORRILL WOLFLE

INTRODUCTION

This paper reports a preliminary experiment concerning the relative effectiveness of various time intervals between the conditioned and the unconditioned stimulus in the formation of a conditioned response. A withdrawal movement of the finger to an electric shock was conditioned to the sound made by an electric hammer.

---

*The experimental work was carried on in the Psychological Laboratory of the University of Washington, under the direction of Dr. E. R. Guthrie.

A few studies of the effect of time interval have already been made. Carr and Freeman (1) conditioned rats to withdraw from a blind alley at the sound of a buzzer. Three time relations were investigated: the buzzer sounded one second before the closed end of the blind alley was reached; one second after; and simultaneously with the reaching of the end of the blind alley. The data showed that presenting the buzzer before the end was reached was most favorable for conditioning; while practically no conditioning occurred when the buzzer was sounded one second after the end was touched. Inasmuch as the buzzer was pressed when the rats were at a given point in the maze, the time intervals necessarily varied with the speed of the rats' movements.

Yarbrough (5) investigated seven time relations in the conditioning of rats. He conditioned a turning (withdrawal) movement, originally caused by electrical stimulation, to the sound of a buzzer. The effectiveness of an interval was measured by the number of trials necessary to secure 90 turning responses in 100 consecutive stimulations. Results were reported on 40 rats. The simultaneous presentation of the stimuli was found to require the least number of trials. The number of trials necessary to secure 90% effectiveness became greater as the interval between the buzzer and the shock was increased. When the buzzer was sounded "immediately" after (no definite time was noted) the shock causing withdrawal had been given, the rats learned to react to the buzzer alone more quickly than when it was presented two seconds before the shock. When the buzzer was presented one second after the shock practically no withdrawal responses were made to it alone.

Pavlov (2) has used the salivary reflex in dogs for investigating the relative value of different time intervals, but has made no systematic exploration of the effects of very small intervals. He states, "A considerable choice is allowed in the establishment of conditioned reflexes as regards the time interval between the beginning of the conditioned stimulus and the moment at which it is reinforced by the unconditioned reflex. This interval can be made very short, 1 to 5 seconds, or even a fraction of a second. . . . Alternatively, the length of time of the isolated action of the conditioned stimulus can be made comparatively long, extending over several minutes. . . . These two types of reflexes are designated respectively simultaneous—or to be more precise, almost simultaneous—reflexes, and delayed reflexes." Under the heading, "simultaneous," then, Pavlov places those intervals between 1 to 5 seconds, the conditioned stimulus being presented first. He states that a conditioned response will never result if the conditioned stimulus *follows* the unconditioned. In the reported experiments on which he based his conclusion, intervals of "some 5 to 10 seconds" and

"10 seconds, 5 seconds, or only a single second" are mentioned. In the data given, therefore, the intervals of less than one second are not mentioned.

Schlosberg (3), using the conditioned patellar reflex, reports that "when the interval between conditioned and unconditioned stimuli was varied between .20 and .44 seconds there was no significant difference in the ease of conditioning; but when this interval was decreased below .11 seconds it was more difficult to form a conditioned reflex." This conclusion was based upon the performance of eight subjects, five of whom had been part of another group using different time intervals.

Method

In the present experiment, the unconditioned stimulus was an electric shock from the current of the secondary coil of an inductorium. The subject rested the end of his middle finger against a small electrode, while his palm lay over a larger one. Watson (4) has described a similar device. In order to preclude the possibility of visual cues, the electric sound hammer which produced the conditioned auditory stimulus was placed directly behind the subject. The timing apparatus and polygraph were in another room. The records obtained for each subject showed the presentation of the unconditioned stimulus (shock), the presentation of the conditioned stimulus (sound), and the withdrawal of the finger.

Ninety students from the elementary psychology class were used as subjects, none of whom had previous information as to the nature of the experiment. Each one was given a demonstration of the correct method of placing his hand and finger against the electrodes, with the instructions to lift his finger when he was shocked, and to replace it after the shock was over. No subject was used on more than one interval.

The time relations investigated, differing by steps of a quarter second, ranged from that with the sound .50 seconds *after* the shock to that with the sound 1.50 seconds *before* the shock. Nine intervals were thus studied: two with the sound after, six with the sound before, and one with the two stimuli simultaneous. These time intervals were regulated by the proper setting of two movable contacts, one for the conditioned and one for the unconditioned stimulus, on a revolving disk, which was driven by a 1/10 H.P., 110-volt, D.C. motor. This disk revolved once in 8.5 seconds. A slight variation was occasionally noted in its speed. The average variation for the longest time interval, 1.5 seconds, was about 45 sigma. With the shorter time intervals it was correspondingly less.

It was found in preliminary experimentation that shocks of very short duration did not cause subjects to withdraw their fingers. Hence the electrical stimulation was increased to 1¾ seconds.

With each subject the auditory stimulus was first given alone. Contrary to Watson (4), who reported that he had never obtained the withdrawal reaction from the bell alone prior to electrical stimulation, some of our subjects withdrew their fingers to the sound. The unfamiliarity of the subjects with electrical stimuli, combined with the instructions which did not mention the sound, doubtless contributed to this effect. In each case practice with the sound alone was continued until no movement resulted from two successive stimulations. Next the shock was given alone with a gradual increase in intensity until the subject reacted.

If at any time during the practice period the subject did not withdraw his finger after electrical stimulation, or if he put his finger back before the shock was over, the amount of current was increased. The shock was constant for each subject in the sense that its intensity was just sufficient to cause a withdrawal movement; however, the absolute intensity varied from subject to subject and, in general, was gradually increased as the practice period became longer.

Three stimulations out of every 20 were with the auditory stimulus alone. These three were irregularly placed, since a preliminary experiment showed that, if the sound occurred alone at regular intervals, the subjects learned the order of stimulation and withdrew their fingers only when they would have been shocked.

As previously mentioned, the timing disk completed a revolution in 8.5 seconds. Hence when the conditioned stimulus was presented .50 seconds before the unconditioned, it was also presented 8 seconds after. To preclude the possibility of conditioned responses occurring because of the time relations of the longer interval, it was varied by turning off the motor for a short time, at least once in every 20 stimulations.

The practice period was approximately 55 minutes long, resulting in about 340 presentations of the conditioned with the unconditioned stimulus, and 60 presentations of the conditioned stimulus alone. The amount of conditioning was expressed in percentage: the number of times the subject responded to the conditioned auditory stimulus divided by the number of times it was presented. Any movement of the finger following the auditory stimulus and occurring before the end of the period in which the shock regularly was given was counted as a conditioned response.

RESULTS

The results are presented in Table 1. Many subjects were never conditioned. That the conditioned stimulus was given alone 15% of the time may account for this fact, since it is known that repetition of the unreinforced conditioned stimulus tends to extinguish the response. But the great variation in percentage of conditioning in each interval is still unaccounted for, as the experimental procedure was constant in this respect.

This experiment supports the general opinion that, with man, experimentally established conditioned "reflexes" are rather unstable. However, when the mean percentage of conditioning for each interval was calculated,

TABLE 1

Percentage of Conditioning with Nine Different Time Intervals between the Unconditioned Stimulus (Shock) and the Conditioned Stimulus (Sound), Using Ten Different Subjects for Each Interval

Number of subject	Interval																	
	U–.50–*C*		*U*–.25–*C*		*U*–0–*C*		*C*–.25–*U*		*C*–.50–*U*		*C*–.75–*U*		*C*–1.00–*U*		*C*–1.25–*U*		*C*–1.50–*U*	
	Sex	%	Sex	%	Sex	%	Sex	%	Sex	%	Sex	%	Sex	%	Sex	%	Sex	%
1	W	2	W	0	—	0	M	0	M	0	M	0	W	0	W	0	M	0
2	W	2	W	2	W	0	W	4	M	2	M	4	W	2	M	2	M	0
3	M	3	W	2	M	3	W	5	M	3	W	4	M	2	W	3	M	0
4	W	5	M	2	M	7	W	15	M	18	M	10	M	4	W	4	W	0
5	W	5	M	3	W	10	W	28	W	20	W	13	W	7	W	7	M	5
6	M	6	W	13	M	16	W	28	W	33	M	13	M	14	M	7	M	9
7	M	6	W	13	—	19	W	46	W	38	M	19	M	18	M	12	W	11
8	W	8	W	15	W	21	W	48	M	68	W	33	W	21	M	12	W	11
9	W	17	M	16	W	24	M	53	W	93	W	57	M	22	W	28	W	11
10	M	47	W	67	W	89	W	62	W	93	W	62	M	46	M	40	W	35
Mean		10		13		19		29		37		22		14		12		8
S.D.		13		19		25		21		34		20		13		12		10
S.D.$_{mean}$		4		6		8		7		11		7		4		4		3
Prob.		67		72		83		73		89		85		63		75		

U=the unconditioned stimulus (shock). *C*=the conditioned stimulus (sound).

If *U* is written first, the shock was first presented. If *C* is written first, the sound was first presented. The interval in seconds between the two stimuli is indicated.

Subjects for each interval are arranged in the table according to percentage of conditioning. *W*=woman. *M*=Man.

Prob.=the probability in 100 chances that the obtained difference was a true difference, computed from the ratio of the difference between means of two adjacent intervals to the standard deviation of that difference.

In the table, the probabilities relate to the column under which they are placed and the next succeeding column.

a definite tendency for this average to vary with the length of the time interval is observable.

The .50-second interval between sound and shock, with a mean of 37% as shown in the table, was most effective for conditioning. The curve of the mean percentage for each interval falls off rather gradually on either side of this maximum.

The interval of .25 seconds with the sound first (*C*–.25–*U*) was next most effective, with a mean of 29%. The simultaneous relationship (*U*–0–*C*) was inferior to those previously mentioned, with its mean of 19%, but superior to the .25-second interval in which the unconditioned stimulus was presented first (*U*–.25–*C*). But instead of the curve dropping immediately to the zero line as soon as the simultaneous point was passed, only a slight decrease (from 19% to 13%) was obtained; and even with the .50-second interval, in which the shock was first presented (*U*–.50–*C*), the mean was 10%.

In the other time direction from the maximal interval the amount of conditioning became less as the interval was lengthened. With the 1.50-second interval with the conditioned stimulus first (*C*–1.50–*U*), the mean was only 8%. In view of this fact it is not probable that the 10% conditioning secured with the .50-second interval, in which the shock was presented first, was due to the much more remote time connection with the previously presented sound. It would mean that an 8-second interval was more effective than a 1.50-second interval.

The conclusion is, therefore, reasonable that in this experiment the amount of conditioning is a function of the time interval (at least between the interval of .50 seconds with the shock first to that of 1.50 seconds with the sound first), with no qualitative difference between the three possible time relations (the unconditioned stimulus presented before, after, or simultaneously with the conditioned stimulus), but rather only a quantitative one. This is contradictory to the general opinion that no conditioning is possible when the conditioned stimulus follows the unconditioned; but, inasmuch as the subjects were trained until they did not react to the sound alone in the preliminary series, since each subject was used on only one interval, and since, even with these precautions, as high a percentage of conditioning was found for the two intervals in which the conditioned stimulus was presented later as for the two intervals of 1.25 seconds and 1.50 seconds in which the conditioned stimulus was presented first, it seems probable that conditioning can occur even though the unconditioned stimulus is presented first. The experiment should be repeated with other intensities, with different conditioned stimuli, and with other unconditioned responses.

The probability that the obtained differences in amount of conditioning between the various time intervals were true differences was computed from the ratio, $\frac{Diff.}{S.D._{diff.}}$ (Table 1). Although the differences between adjacent intervals are not statistically significant, the general shape of the curve and the lack of irregularities in it make quite probable the conclusion that a repetition of the experiment would disclose approximately the same relative effectiveness of time intervals.

SUMMARY

In conditioning a withdrawal movement of the finger to an auditory stimulus, the number of conditioned responses in a 55-minute practice period was influenced by the time intervals between stimuli.

With an interval series in steps of .25 seconds, beginning with the unconditioned stimulus (shock) .50 seconds *before* the conditioned stimulus (sound), and ending with the shock 1.50 seconds *after* the sound, the mean percentage of conditioning for the separate intervals was 10, 13, 19, 29, 37, 22, 14, 12, and 8.

The maximal amount of conditioning was 37%, using the .50-second interval with the sound presented first, which, therefore, in this experiment represents the optimum interval for conditioning.

Even when the sound was presented .50 seconds after the shock, the conditioning amounts to 10%.

The results of this experiment are not in fundamental disagreement with other investigations. Yarbrough also found conditioning when the conditioned stimulus was given later; Pavlov has not investigated the effect of very small intervals; Schlosberg found the time range between .20 and .44 seconds more effective than other relationships.

REFERENCES

1. CARR, H. A., & FREEMAN, A. S. Time relationships in the formation of associations. *Psychol. Rev.*, 1919, **26**, 465-474.
2. PAVLOV, I. P. Conditioned reflexes: an investigation of physiological activity of the cerebral cortex. (Trans. by G. V. Anrep.) London: Oxford Univ. Press, 1927. Pp. xv+430.
3. SCHLOSBERG, H. A study of the conditioned patellar reflex. *J. Exper. Psychol.*, 1928, **11**, 468-494.
4. WATSON, J. B. Psychology from the standpoint of a behaviorist. Philadelphia: Lippincott, 1919. Pp. ix+429.
5. YARBROUGH, J. U. The influence of the time interval upon the rate of learning in the white rat. *Psychol. Monog.*, 1921, **30**, No. 135. Pp. 52.

Ohio State University
Columbus, Ohio

Psychom

A PRELIMINARY STUDY OF OCULAR ATTITUDES IN THINKING OF SPATIAL RELATIONS

Edward G. Stoy

The Problems

This study was concerned with three general problems. The first was the determination of the relative size and frequency of horizontal eye-movements in thinking about spatial as compared with non-spatial relations. A second problem was the determination of the existence of individual differences in the size and frequency of eye-movements in thought. The third problem was concerned with a determination of the existence of relationship between size and frequency of eye-movements and ability to deal with spatial concepts.

Historical

This study was undertaken as a result of the writer's interest in the more general problem of the rôle played by motor attitudes in thinking of objects. It was deemed advisable to restrict the present investigation to a special aspect of this general problem, viz., the rôle played by *ocular* attitudes in thinking of *spatial* distinctions and relations.

So far as the writer is aware, no investigations of his particular problem have been reported. However, a number of discussions of the general problem have been published. Carr (1) has covered it briefly, as follows (p. 174 ff.):

> "Any perceptual reaction to an object includes some attentive adjustment of the sense organs, and these motor activities are frequently aroused in thought and serve to signify in part the sensuous character of the object. A vocal process is an essential component of nearly all of the author's representative devices, but in addition to this element, he finds it necessary to employ an ocular adjustment when he desires to think of the visual aspect of a bell and an auditory adjustment to apprehend it as a sounding object. These sense organ adjustments are also frequently employed as a means of thinking of the location of things. Let one imagine a visual object at a distance of fifty feet, and then suddenly think of it as located within a few feet of the face. A very pronounced convergent strain can usually be detected as a result of this effort. Likewise some appropriate adjustment of the body, head, or eyes can usually be observed when one quickly imagines an object as located to his extreme right or at some distance directly over the head.
>
> "The well-known phenomena of muscle reading is based upon the fact that a thought of the location of an object naturally involves some appropriate adjustment of the organism in reference to that position....Many people are prone to use various facial expressions and gestures in expressing their thoughts to others, and naturally such people may employ these motor activities in thinking for themselves. Many people are able to think of yes and no by means of slight movements of the head. In fact, language might well be defined as a device for thinking publicly—to express thoughts to others, and naturally there is no reason why the same devices should

not be employed in thinking privately. A young lady, who had occasion to use the deaf-mute language quite frequently since early childhood, once informed the author that a considerable portion of her thinking was carried on in terms of slight movements of her fingers. As a final example of the devious ways of thinking, witness the extent to which children use their fingers as symbols in numerical computations. Practically any item of experience is capable of being used in the service of thought."

The present investigation is closely allied with a number of studies dealing with other aspects of the more general problem. For example, a quite similar study is that of Perky (3), who investigated the rôle of ocular movements and attitudes in distinguishing between images of memory and images of imagination, or, in functional terms, between imaginary objects and those of memory.

This study is less closely connected with a number of experiments designed to test the linguistic or laryngeal theory of thought. Examples of these are the studies made by Lashley and by Wyczoikowska (5). Watson (4) recorded Lashley's work as follows:

"With a delicate apparatus which recorded the tongue movements in two dimensions he was enabled to show that the overt but whispered repetition of a sentence produced a tracing on the smoked drum which was wholly similar except for amplitude to that obtained when he told the subject to think the same thing without making overt movements. He was enabled to verify this again and again."

Wyczoikowska obtained records of tongue movements on a kymograph drum during thought. These records were supplemented by introspective material in which the subjects stated that feelings of movement were present.

General Procedure

The 50 subjects in these experiments were first-year students at Tilden Technical High School, Chicago.

Each student was tested individually and was asked to solve 11 spatial and 10 non-spatial problems of the types shown below.

Spatial Problems

1) Imagine that you are jumping up a hill. Every time you jump up 3 feet you slide back 2 feet. The hill is 30 feet long. In how many jumps can you reach the top?

2) Imagine that you are on top of a building at State and Madison Streets. You can see the whole city. How much of Chicago (what fraction) would you see west of State Street?

3) Think of a glass of water over on the desk. There is a 2″ diameter ball floating on the water, half-submerged. The top of the ball is level with the top of the glass. How far below the top of the glass is the level of the water?

Non-Spatial Problems

1) All kodaks are cameras, but all cameras are not kodaks. Is this true or false? Why?

2) In your opinion what is the best and most important effect of prohibition?

3) No news is good news; therefore all news is bad news. Is this true or false? Why?

The selection of these problems as being spatial or non-spatial, of course, was more or less arbitrary. However, there was good agreement among competent judges on the distribution.

The problems were stated orally to the subject, who was seated about 8 feet from a blank wall, at which he was instructed to look. During the time that the subject solved the problem, eye-movements were observed with the aid of a Bausch and Lomb keratometer. The keratometer is a tubular instrument about 12 inches long and 1 inch in diameter whose lens equipment gives a magnification of about one and one-half diameters. At one end it contains a scale ruled in half millimeters. A measure of horizontal eye-movements was obtained by observing the excursion of the outside edge of the iris across this scale. For this reading the keratometer was held at an angle of 45° to the subject's line of vision, and it was pressed lightly against his brow so that it would remain stationary.

Measurements of horizontal eye-movements were recorded from the time the experimenter started to repeat the problem to the time at which the subject finished his answer. A record of this length of time was obtained with a stop-watch. In addition to counting on the Veeder counter the number of horizontal movements, measured to the nearest millimeter, the experimenter kept in mind the largest single excursion which occurred during the solution of the problem. Also, a brief notation on the subject's answer was made.

Thus, for each of 50 subjects there were obtained measures of largest single movement, the number of movements, and the time on each of 21 problems—11 spatial and 10 non-spatial.

RESULTS

The first question to be answered by the results of this study was that concerned with the relative size and frequency of movement during thought on spatial as compared with non-spatial problems. That is, do larger single movements occur, on the average, in the solution of spatial problems, and are the measurable movements more frequent?

It was possible to translate the data obtained during the experiment into average largest movement made by each subject on the 11 spatial problems and into the same measure for the 10 non-spatial problems. Likewise, the average number of movements per second made by each subject on the 11 spatial problems was calculated, and also the same measure for the 10 non-spatial problems.

When the difference between the mean of 50 averages of largest movement on spatial problems and the mean of 50 averages of largest movement on non-spatial problems was divided by the probable error of the difference, the figure 7.535 was obtained. The same procedure on frequency yielded a significance ratio of 6.318. Since we were not concerned with the difference in group performances on one activity, but rather with the difference in average performance of the whole group on two groups of activities, the probable errors of difference referred to above were calculated from $P.E._{diff.} = \sqrt{P.E._{\cdot 1}^2 + P.E._{\cdot 2}^2 - 2r_{12}\, P.E._{\cdot 1}\, P.E._{\cdot 2}}$, where r_{12} is the coefficient of correlation between 50 average largest movements on spatial and non-spatial, or the coefficient of correlation between the 50 average movements per second on the two sets of problems. These coefficients are, respectively, .654±.055 and .718±.049.

It may be noted from the table that the two ratios of 7.535 and 6.318 are statistically significant, in that they indicate almost certainly that in a similar repeated experiment the mean average measures for spatial problems would be greater than these measures for non-spatial problems.

It may be noted also that both the correlation coefficients are significant and indicate that, on the whole, those who show large eye-movements (as measured by the largest single movement) on spatial problems also show large movements on non-spatial problems, and that those who show large frequencies of movement on one type of problem show large frequencies also on the other type.

The third general problem was the determination of the relation between individual differences in eye-movement and individual differences in ability to deal with spatial concepts. Two different types of criteria of ability to deal with spatial concepts were used in these determinations, namely, ability shown in working the problems used in this study and ability to learn architectural drawing. Each of these two types of criteria in turn yielded two separate measures; ability to work the problems was measured by correctness of solution and time required to obtain a correct solution, and ability to learn architectural drawing was measured both by class grades and by instructors' estimates.

The first criterion of "ability to deal with spatial concepts" was the number of problems solved correctly, but this was hardly satisfactory, since few of the problems carried definite answers. When the number of problems right (of these eight) was correlated with average largest movement, a coefficient of —.049±.095 was obtained. A correlation between the same criterion and the average number of movements per second yielded a coefficient of —.026±.095.

The second criterion of ability to deal with spatial problems was the time required for a correct solution. It was desired to correlate the two measures of eye-movement with this criterion. However, only five of the eight spatial problems carrying definite answers were solved correctly by 20 or more of the 50 subjects, and the populations, even so, ranged from 22 to 31. Although the correlation between the criterion and the largest movement was straightforward, a correlation between the criterion and the second measure of eye-movement (movements per second) would, in all probability, have been spurious. Also, a correlation between the number of movements and the criterion would have been questionable. Consequently, correlations were calculated only between the criterion and the largest movement. Of five such correlations, only one was significant ($r = .511 \pm .113$). There is some indication here, then, that those who showed the largest movements also took the longest time in solving the problem; or, in other words, those who showed the largest movements had the least ability, as measured by this criterion.

Since the difficulties involved in the third criterion of "ability to deal with spatial concepts" (namely, class grades in architectural drawing) were many, and since few valid results were obtained, it may be best to consider now the results of the use of our last criterion, namely, instructors' estimates of ability to learn architectural drawing.

The instructors of the two classes of which the subjects of this study were members were asked to place their students in rank order as to ability to learn architectural drawing. They were asked specifically to partial out, as far as possible, such factors as perseverance, tardiness, and conduct in class. Only "ability to catch on" was to be considered. Class grades were not consulted at the time the ranking was done. The ranking was done twice by each instructor, the intervening time being one week. At the time the first ranking was made the instructors were not informed that the work was to be repeated.

Instructor A had a class of 30 students, and the reliability of his ranking was .977, as measured by correlating the two sets of ranks, made one week apart. Instructor B showed a reliability coefficient of .896 measured in the same manner, on 20 students.

The instructor's rank on each student was in these two groups transmuted into a percentile rank, and from this into a sigma value. Each group, of course, was treated separately.

When the measures of ocular attitudes on each problem were correlated by the Pearson method with the sigma values of instructors' judgments, coefficients were obtained as shown in Table 1. In addition to the figures for separate problems, there will be found at the bottom of the table the coefficients yielded by correlating the averages of the eye-movements shown on all the problems against the criterion.

It will be noted that five of the coefficients are .25 or over, and are marked with an asterisk. If a ratio of three between a coefficient and its probable error is arbitrarily accepted to denote significance, these marked coefficients are significant, for the probable error of a coefficient of .250 based on 50 cases is .089.

TABLE 1

Coefficients of Correlation between Criterion of Ability to Deal with Spatial Concepts and Measures of Ocular Attitude

Spatial problem number	Criterion *vs.* largest movement	Criterion *vs.* movements per second	Non-spatial problem number	Criterion *vs.* largest movement	Criterion *vs.* movements per second
1	.188	—.135	1	.245*	.064
2	.042	—.069	2	.019	.041
3	—.032	—.253*	3	.207	.206
4	.354*	—.150	4	—.136	—.155
5	.000	.112	5	.052	.249*
6	—.216	—.019	6	.000	—.086
7	.030	—.026	7	.117	.085
8	.155	.068	8	—.106	.016
9	.159	.051	9	.250*	.220
10	.097	—.057	10	.021	.084
11	—.111	.008	Av.	.025	.206
Av.	.111	.000			

One of these five coefficients indicates the correlation between a measure of frequency of movement on a spatial problem and the criterion, another is concerned with a measure of size of movement on a spatial problem, another with frequency on a non-spatial problem, and two with size on non-spatial problems. It is worthy of note that the first-mentioned coefficient is negative, that is, the lower the frequency of movement on this problem the greater the ability to deal with spatial concepts, as measured by the criterion.

These diversified problems indicated a possibility of using a larger number of similar problems for predicting ability to learn to deal with spatial concepts, as this ability was measured by the present criterion. A larger number of problems in the examination should increase the reliability of the measure when the whole battery of problems is considered.

In order to determine the optimal predictive value of the group of five problems as they now stand, a multiple R was calculated. This correlation yielded a multiple coefficient of $R = .600 \pm .061$.

SUMMARY

1) The mean average largest horizontal eye-movement was found to be greater when spatial problems were solved than when non-spatial problems were solved. That is, the mean of 50 average largest movements on the former material (11 problems) was 2.432 mm. The mean average largest movement on 10 non-spatial problems was 1.732 mm. This difference of means amounted to 7.535 times its probable error, and was thus significant.

The mean average frequency of horizontal eye-movement was greater when spatial problems were solved than when non-spatial problems were solved. The mean of 50 averages of number of movements per second was 0.477 for 11 spatial problems and 0.326 for 10 non-spatial problems. This difference of means amounted to 6.318 times its probable error.

2) Individuals exhibited differences in the size and frequency of horizontal eye-movements during thought. Those who showed large movements in thinking of spatial problems also showed relatively large movements in thinking of non-spatial problems. This was indicated by a correlation coefficient of .654 between average measures of largest movements on spatial problems and the same measures on non-spatial problems.

That there were individual differences in frequency was shown by a correlation coefficient of .718 between average number of movements per second on spatial problems and the same measures on non-spatial problems. These coefficients of .654 and .718 are significant in high degree.

3) There is little evidence that eye-movement is correlated with ability when success is measured by ability to answer problems correctly or to answer them quickly. Also, there is little evidence of such a relationship when ability is measured by class grades in architectural drawing. However, there is evidence that eye-movements are indicative of ability to deal with spatial concepts when the latter is measured by instructors' judgments. Five correlation coefficients of .25 or over were obtained between measures of size and frequency of eye-movement and instructors' judgment of ability to learn architectural drawing. When properly weighted and combined these five individual problems will yield a coefficient of .600±.061.

REFERENCES

1. CARR, H. A. Psychology. New York: Longmans, Green, 1925. Pp. 432.
2. KELLEY, T. L. Statistical method. New York: Macmillan, 1923. Pp. xi+390.
3. PERKY, C. W. An experimental study of imagination. *Amer. J. Psychol.,* 1910, **21**, 422-452.
4. WATSON, J. B. Psychology from the standpoint of a behaviorist. Philadelphia: Lippincott, 1919. Pp. ix+429.
5. WYCZOIKOWSKA, A. Theoretical and experimental studies in the mechanism of speech. *Psychol. Rev.,* 1913, **20**, 448-458.

University of Chicago
Chicago, Illinois

AN EXPERIMENTAL STUDY IN PERCEPTUAL INSIGHT

GEORGE F. DUNKELBERGER AND ELDON K. RUMBERGER

How much of our "image-life" is governed by insight? This question, while directly stated, can by no means be answered as directly. This matter of insight has become a sort of psychological football which is being tossed back and forth between opposing sides. We can hardly say that it has come to rest in any position from which it will not sooner or later be dislodged. In an effort to throw some little light upon this particular phase of Gestalt psychology, the following experiment was conducted. We wish it to be clearly understood that this problem was approached with an open mind. We do not aim to disprove any one phase of Gestalt but rather to ascertain to what extent it really functions. Furthermore, we wish to advise that this is but the beginning of a series of experiments along this line.

It was noted that, in Köhler's experiments with apes, the animals showed a marked deficiency in the discrimination of forms. It is true that these animals were able to solve problems which involved the use of two sticks joined together, but that they were unable to utilize boxes to the best advantage to reach a desired goal. It was this latter fact that inspired this experiment. While we did not work with apes, we feel sure that a problem which cannot be solved by the human mind is certainly beyond the comprehension of the chimpanzee.

The materials used in this experiment consisted of two sets of blocks and a base upon which they could be built to just reach two horizontal bars (for full description see Figure 1). These blocks were so made that they would stand but one way and reach the bar for which they were intended. None other but the one arrangement would solve the problem.

Preliminary trials were made with elementary and high school pupils with the result that they were unable to solve the problem except by trial and error. It was decided to start the experiment with college freshmen. We worked upon the assumption that if individuals of this type were unable to solve the problem by insight we could not expect a solution from subjects of a lower educational class.

The procedure was carried out as follows. A subject upon entering the room was informed that he should be at his ease so far as this test was concerned. We made it clear to the subject that it was not an intelligence test. The subject was seated at a table on which the materials described were already placed. An instruction sheet was given him which explained fully what he was to do with the blocks. If he had any questions that could be answered without disclosing the solution, they were answered.

Forty subjects were tested as to their ability to solve the problem by insight. They were permitted to handle the blocks as much as they pleased, but they were not allowed to build with them until they had consummated their image plan for the correct arrangement of the blocks of the particular set. All subjects began with Set A and proceeded to Set B after they had solved or failed to solve Set A. Each subject of this first group was allowed but one trial in building with each set. The subject was timed during the interval of handling the blocks and building until stopped by the operator in case he made an error. The data in each case was recorded as shown in

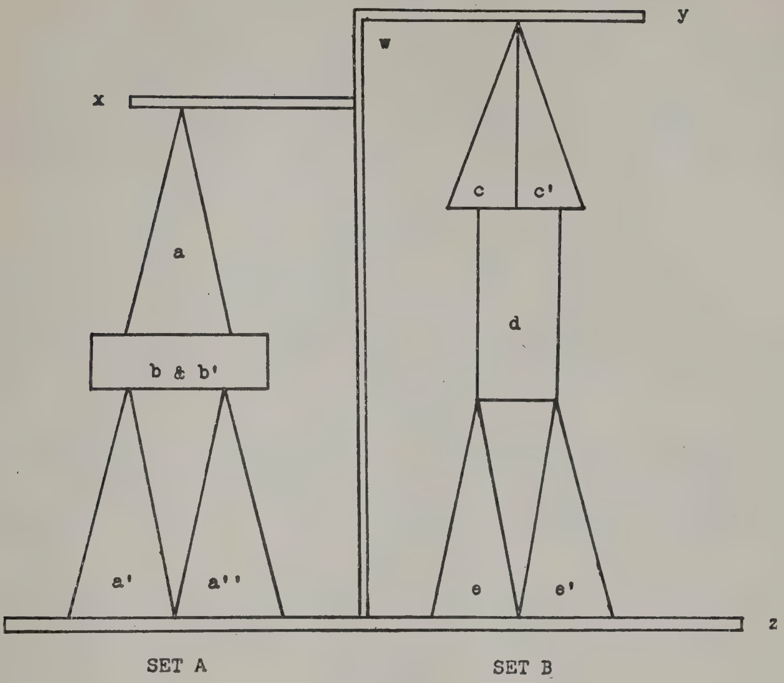


FIGURE 1

This figure represents the blocks placed in their proper relationship in the correct and only solution of the problem. *a, a'* and *a''*, represent the three identical triangular blocks (isosceles). *b* and *b'* represent two right triangular blocks placed together to form a rectangular block with the right triangular surfaces resting upon the apex of blocks *a'* and *a''*. *c* and *c'* represent two right triangular blocks. *d* represents a rectangular block. *e* and *e'* represent two isosceles triangular blocks with apexes and bases slightly planed. *z* represents the base upon which the blocks were built to reach the horizontal bars *x* and *y* which extend from the column *w*, supported by the base *z*.

Table 1. The columns headed *S* or *F* record solutions or failures. All other columns are self-explanatory.

A second group of 40 subjects was tested as to their ability to solve the problem by trial and error. The procedure was the same as for the first group with the exception that the subjects were permitted to build the blocks as many times as they wished until they solved the problem or announced their inability to do so. They were timed from the moment they began handling the blocks. The data was recorded as above.

The subjects in this experiment ranged in age from 17 to 23 years. The

TABLE 1

No.	Sex	Age	Educational group	*S* or *F* Set A	*S* or *F* Set B	Time Set A	Time Set B
1.	M	19	freshman	F	F	3:14	4:22
2.	F	17	freshman	S	F	7:56	5:42

TABLE 2

Results of Experiment

Insight—Set A		
Average time of manipulation by	34 M subjects	2 min., 1.53 sec.
Average time of manipulation by	6 F subjects	2 min., 15.00 sec.
Average time of manipulation by	40 subjects	2 min., 6.5 sec.
Percentage of solution by 40 subjects		0
Insight—Set B		
Average time of manipulation by	34 M subjects	1 min., 37.7 sec.
Average time of manipulation by	6 F subjects	1 min., 38.7 sec.
Average time of manipulation by	40 subjects	1 min., 37.9 sec.
Percentage of solution by 40 subjects		5.0
Percentage of solution by 34 M subjects		5.8
Percentage of solution by 6 F subjects		0.0
Trial and Error—Set A		
Average time of manipulation by	27 M subjects	4 min., 14.7 sec.
Average time of manipulation by	13 F subjects	5 min., 25.3 sec.
Average time of manipulation by	40 subjects	4 min., 37.6 sec.
Percentage of solution by 40 subjects		47.5
Percentage of solution by 27 M subjects		55.5
Percentage of solution by 13 F subjects		30.8
Trial and Error—Set B		
Average time of manipulation by	27 M subjects	3 min., 57.5 sec.
Average time of manipulation by	13 F subjects	4 min., 14.7 sec.
Average time of manipulation by	40 subjects	4 min., 3.2 sec.
Percentage of solution by 40 subjects		62.5
Percentage of solution by 27 M subjects		74.1
Percentage of solution by 13 F subjects		38.4

average age was 19.2, and the mode was 19. As for the results of this experiment, let us say that it can be briefly summed up in the preceding tabular form (Table 2).

Three months later 70 members of the incoming freshman class were tested in the manner above described. Thirty-five were tested using the insight method and 35 by trial and error.

The subjects in this group ranged in age from 16 to 22 years. The average age was 18.18 years and the mode was 18.

The results of this second experiment appear in like tabular form as shown in Table 3.

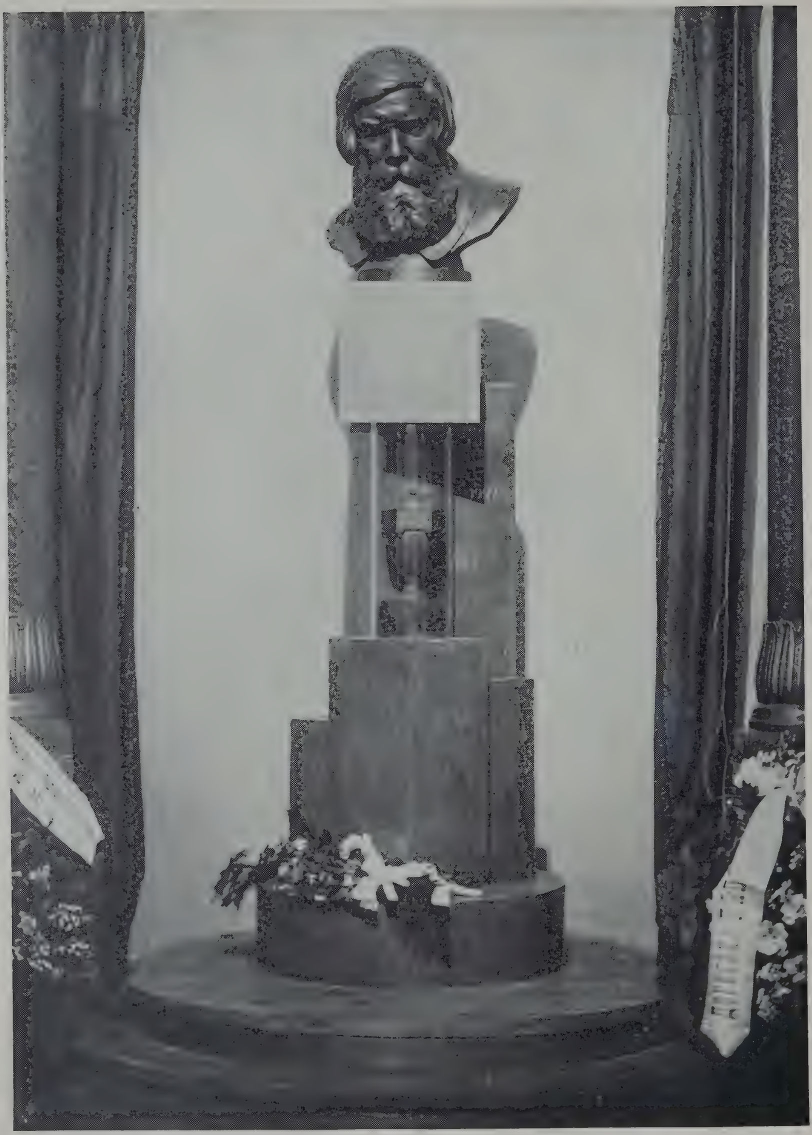
TABLE 3

Insight—Set A				
Average time of manipulation by	23	M	subjects	2 min., 54.27 sec.
Average time of manipulation by	12	F	subjects	3 min., 17.08 sec.
Average time of manipulation by	35		subjects	3 min., 2.09 sec.
Percentage of solution by	23	M	subjects	4.34
Percentage of solution by	12	F	subjects	0.00
Percentage of solution by	35		subjects	2.85
Insight—Set B				
Average time of manipulation by	23	M	subjects	2 min., 15.96 sec.
Average time of manipulation by	12	F	subjects	1 min., 56.67 sec.
Average time of manipulation by	35		subjects	2 min., 9.46 sec.
Percentage of solution by	23	M	subjects	4.34
Percentage of solution by	12	F	subjects	0.00
Percentage of solution by	35		subjects	2.85
Trial and Error—Set A				
Average time of manipulation by	27	M	subjects	8 min., 49.44 sec.
Average time of manipulation by	8	F	subjects	12 min., 39.12 sec.
Average time of manipulation by	35		subjects	9 min., 42.23 sec.
Percentage of solution by	27	M	subjects	66.67
Percentage of solution by	8	F	subjects	75.00
Percentage of solution by	35		subjects	68.57
Trial and Error—Set B				
Average time of manipulation by	27	M	subjects	4 min., 4.43 sec.
Average time of manipulation by	8	F	subjects	10 min., 52.37 sec.
Average time of manipulation by	35		subjects	5 min., 37.82 sec.
Percentage of solution by	27	M	subjects	85.18
Percentage of solution by	8	F	subjects	62.5
Percentage of solution by	35		subjects	80.0

We hesitate to draw any final conclusions, but prefer to let the results, as above tabulated, speak for themselves. This study constitutes the beginning of a series of investigations in the field. We would, however, make this rather noteworthy statement. It was noticed in the trial-and-error experiments that the subjects in every case repeated errors several times in arranging the blocks. This in itself seems to point to the fact that the degree of insight involved is questionable.

After we have carried out further work in this line with larger and more varied groups, the results may throw an entirely different light upon this problem. The results as they stand here would indicate that individuals of this particular class and age display very little insight in the matter of perception. Their form-life and image-life seem to be directed chiefly through the medium of trial and error.

Susquehanna University
Selinsgrove, Pennsylvania



THE MONUMENT OF V. M. BEKHTEREV

According to Bekhterev's wish, his ashes have been placed in the Museum of the Reflexological Institute for Brain Researches. The urn containing the ashes is enclosed in a glass cylinder at the top of which is placed a bust of Bekhterev. Bekhterev's schematic outline of the conduction paths is engraved on one of the faces of the monument.

SECOND ANNIVERSARY OF THE DEATH OF BEKHTEREV

A. L. Schniermann

On the 24th of December, 1929, the Bekhterev Reflexological State Institute for Brain Researches commemorated the second anniversary of the death of Professor V. M. Bekhterev, the founder and former director of the Institute.

Vladimir M. Bekhterev was born in 1857 and died in 1927. He was the author of some six hundred scientific papers and books in the fields of anatomy, physiology, psychology, reflexology, pedology, defectology, and psychiatry. He created a new tendency in psychology, i.e., Human Reflexology. The chief works of Professor Bekhterev are *Conduction Paths of the Brain and Spinal Cord* (1888), *Bases of the Teaching Concerning the Function of the Brain* (1903-1907), *Objective Psychology* (1907-1912), and *General Principles of Human Reflexology* (1918). Some of his studies which were not published during his life will appear soon.

A scientist of very broad conception, Bekhterev was also an excellent teacher, and many of his former students are now professors in the U. S. S. R. or in other countries. Endowed with a remarkable ability for organization, he founded a number of institutions for the study of human behavior, normal and abnormal, child and adult.

In the history of Russian psychoneurology the name of V. M. Bekhterev occupies a very high position.

Bekhterev Reflexological State Institute
Leningrad, U. S. S. R.

THE INFLUENCE OF TWO DIFFERENT INTERPOLATIONS UPON TIME ESTIMATION

Paul L. Whitely and J. Carver Anderson

Introduction

In the experimental investigation of temporal judgments it has been repeatedly disclosed that "filled" time is judged to be shorter than "unfilled." Casual observation is in agreement with this disclosure. From casual observation, the writers have been impressed by, what we may call for convenience, a rhythmic factor in the perception of time. For example, at a vaudeville show, one gains the impression that the playing of music between acts makes the duration seem shorter. Furthermore, there comes the suggestion that the particular character of the rhythm of the music may be a conditioning factor of time estimation. The present investigation was designed to test the influence of two different interpolations upon the estimation of time, namely, that of a non-rhythmical buzzer tone and that of music of a decidedly rhythmical character. The buzzer tone was provided by means of an ordinary office buzzer, while the music consisted of an Edison record, the *Radetsky March* by J. Strauss, as played by the United States Marine Band.

Two experiments were performed. In the first experiment, three temporal durations—15 seconds, 30 seconds, and 45 seconds—were used, each of which was judged under three different conditions, namely, (1) when there was no filling of the temporal duration, (2) when the duration was filled by the buzzer tone, and (3) when music filled the temporal duration. The subjects, who were college students, were instructed to estimate the time that had elapsed in terms of seconds and to report to the experimenter verbally. The method of paired comparisons was used in the second experiment. The first duration was taken as standard and the second was judged as equal to, shorter than, or longer than the standard. In this experiment, 5-second, 10-second, and 15-second durations were employed. Further details concerning the procedure can best be explained in connection with the discussion of each experiment separately.

Description of Apparatus

In order to give the necessary fillings for the various durations and to insure as great accuracy as possible, all signals were given by means of electrical devices. The contacts were made by standard telegraph keys, so closely adjusted as to eliminate all loose play. An ordinary buzzer for office use, selected for constancy of tone, was employed to give the nonrhythmical buzzer tone. A standard telegraph sounder gave the clicks to bound the unfilled intervals. Since the standard sounder gives two clicks when a contact is made, one when the sounder arm is pulled down by the magnet and a second when it is released, a small strip of soft rubber was fixed on top of the arm, so that but one click was made. Dry cells were used to furnish the necessary current. These were frequently replaced in order to keep constant the volume of the tone and the intensity of the click.

Figure 1 is a diagram of the side view of the apparatus designed to lower and raise the needle of the phonograph. In the figure, the dotted lines represent the needle, the diaphragm, and the removable tone arm of the standard Edison Laboratory model phonograph. The solid lines show the necessary parts, fitted in place. The needle *J* is fastened to the underside of the holder *A*, which was fixed on a piece of spring steel and connected to the tone arm at *K*. The diaphragm is inside at *B* and the whole arm is fixed to the end of the amplifying arm at *L*. A strip of metal, *C*, was bent in the shape shown, and screwed to *A*. This strip and *A* were thus connected so that when *C* was raised, *A* was raised, and the needle also; thus the music stopped. When *C* and *A* were lowered, the needle again touched the record, and the music began. To the main part of the tone arm, the bracket *F* was fixed. This bore, on its top, a strip of metal, to the other end of which the magnet *D* was fastened.

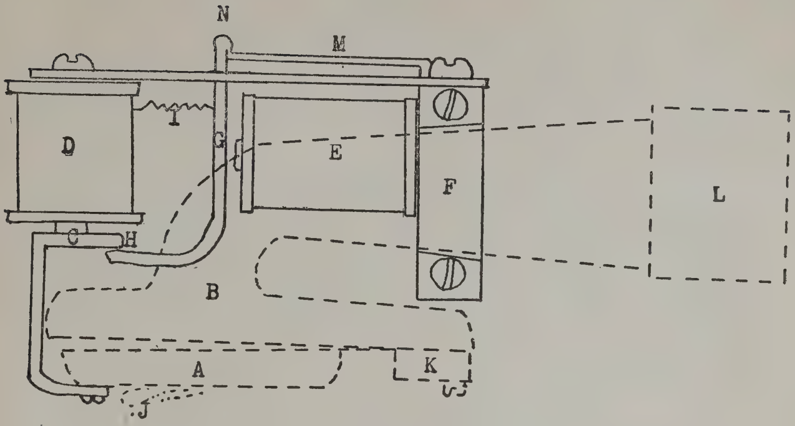


FIGURE 1
Diagram of Apparatus

When a current was passed through this magnet *D*, the strip *C* and the needle were raised. There were also connected to *F*, the magnet *E* and the strip *M*. *M* bore a supporting arm *G*, on one end, connected to *M* by the hinge *N*. Spring *I* contracted when *C* was raised and pulled the supporting arm *G* in its direction. Magnet *E* pulled the arm *G* when a current was passed through it. When there was no music desired, supporting arm *G* held *C* up so that the needle was not in contact with the record.

The method of operation was as follows: To start the music, telegraph key No. 1 was closed, magnet *E* pulled strip *G*, allowing *C* to drop. While the music continued, *G* rested on the end of *C*, at *H*, so that no current was necessary to hold it back. To stop the music, key No. 2 was closed, magnet *D* pulled *C* up, spring *I* pulled arm *G* under *C*, and when the key was released, *G* continued to support *C*. The phonograph record was allowed to rotate constantly, so that the operator had merely to press the telegraph keys to give the necessary signals.

First Experiment

As indicated above, the subjects were instructed in the first experiment to observe the duration under three different conditions, and to report to the experimenter in seconds the length of time which he estimated the duration to be. The procedure can best be explained by reference to an illustrative cycle, which is represented in Table 1. An explanation of the symbols (left to right) is as follows: *N* means that the duration was unfilled; *T* signifies that a non-rhythmical buzzer tone filled the duration; while *R* indicates that rhythm or music filled the duration. The three different dur-

TABLE 1

ILLUSTRATIVE CYCLE

*N*15	*T*30	*R*45	*N*30	*T*45	*R*15	*N*45	*T*15	*R*30
*T*30	*R*45	*N*15	*T*45	*R*15	*N*30	*T*15	*R*30	*N*45
*R*45	*N*15	*T*30	*R*15	*N*30	*T*45	*R*30	*N*45	*T*15

ations used are indicated by the numbers following the letters. The procedure may be clarified more definitely by way of illustration. The first subject began with the condition *N*15 and followed the order in the illustrative cycle. Subject number two began with the *T*30 condition and ended with the *N*15 condition. The third subject began with *R*45 and ended with *T*30, and so on. Twenty-seven subjects participated in the experiment, and each of the subjects went through the cycle three times on successive days. Thus, it will be observed, the subjects were put through the various conditions in a counterbalanced order so that all conditions had equal average position. Each subject made nine judgments for each experimental condition; in all, there were 243 judgments per condition.

In Table 2 are to be found the results of the first experiment. In the first column are shown the various conditions employed. There follow, in the second and third columns, the medians and the means for the various conditions. The numbers in the fourth column represent the deviation of the mean from the actual duration. The percentage of accuracy of judgments and the σ's of the distribution are to be found in the fifth and last columns, respectively.

An inspection of the tabular data reveals the fact that there is perfect consistency of the mean scores for the three durations of 15 seconds, 30 seconds, and 45 seconds. In each case, the *R* condition is judged to be shortest, whereas the *N* condition is judged to be longest. The mean scores for the *R*15, *R*30, and *R*45 conditions are 13.03, 24.36, and 35.89, respectively, as compared with 14.52, 26.10, and 38.35 for the *N* conditions. Also, the median scores for the *R* conditions are consistently smaller. The difference between the median scores for the *N* and *T* conditions for the 15-second and 30-second durations are very slight, as are also the means for the 30-second and 45-second durations.

These same results may be interpreted from the standpoint of the accuracy of the judgments. From the figures in Column 5, we observe that the *N* condition makes for greatest accuracy of judgment, whereas the *R* condition renders least accuracy of judgment. The results are indicated in terms of percentage of accuracy.

The statistical reliability of the data is shown in Table 2*a*, which includes the difference of the means and the σ's of the difference. In no instances is a difference to be found which exceeds its σ by three times, and therefore the differences are not reliable according to this statistical criterion. The greatest reliability is indicated in the case of the *N* and *R*

TABLE 2

RESULTS OF THE FIRST EXPERIMENT

Condition	Median	Mean	Dev. from Act. Dur.	Percentage Accuracy	$\sigma_{dist.}$
*N*15	13.16	14.52	.48	.97	6.99
*T*15	13.36	13.72	1.28	.92	6.93
*R*15	12.66	13.03	1.97	.87	6.24
*N*30	23.33	26.10	3.90	.87	12.55
*T*30	23.13	25.78	4.22	.86	11.85
*R*30	21.73	24.36	5.64	.81	12.75
*N*45	33.05	38.35	6.25	.86	19.10
*T*45	34.90	38.27	6.73	.85	18.80
*R*45	32.49	35.89	9.11	.80	17.04

TABLE 2*a*

DIFFERENCES OF MEANS AND SIGMAS OF THE DIFFERENCE FOR THE DATA IN TABLE 2

Conditions compared	Difference in Means	$\sigma_{diff.}$
*N*15-*T*15	.80	.63
*N*15-*R*15	1.49	.59
*T*15-*R*15	.69	.59
*N*30-*T*30	.32	1.10
*N*30-*R*30	1.74	1.14
*T*30-*R*30	1.42	1.12
*N*45-*T*45	.08	1.71
*N*45-*R*45	2.46	1.64
*T*45-*R*45	2.38	1.62

comparisons for the three durations. While the data do not show reliability according to the above mentioned criterion yet the consistency of the results does suggest that the *R* and *T* conditions may be somewhat influential in affecting temporal judgments, and of these two conditions, the *R* condition, or music, is more effective than the non-rhythmical buzzer tone.

An examination of the raw data suggests that some of the subjects were differently influenced by the experimental conditions; as the experiment was not controlled for the purpose of studying individual differences, however, no definite statements can be made regarding them.

SECOND EXPERIMENT

Again, the procedure employed in the second experiment may best be explained by reference to an illustrative cycle, which is shown in Table 3. The symbols *N, R,* and *T* refer to the three experimental conditions pre-

TABLE 3

Illustrative Cycle

*NR*5	*NT*10	*TR*15	*NT*5	*TR*10	*RN*15
*TR*5	*RN*10	*TN*15	*RN*5	*TN*10	*RT*15
*TN*5	*RT*10	*NR*15	*RT*5	*NR*10	*NT*15

viously explained. The numbers indicate, in terms of seconds, the three different durations used. The subject was instructed that two durations would be presented in succession. The first duration was to be taken as standard, and the second was to be judged as equal to, shorter than, or longer than, the standard. Actually, the two durations presented for comparison were always the same, but they occurred under different experimental conditions. For example, the first subject began with the *NR*5 condition. A duration of five seconds, unfilled, was immediately followed by a five-second duration which was filled with music. The interval between the two durations was constant for each pair. The subject then made a judgment, as indicated above, after which the next pair, *NT*10, was presented, and so on from left to right. The second subject began with the *NT*10 condition. Eighteen subjects were necessary to complete the cycle in a counterbalanced practice series, and the cycle was completed by a second group of 18 subjects. Seventy-two judgments were made for each duration and experimental condition, or 216 judgments for each experimental condition, including all three durations.

The data for this experiment are to be found in Table 4. In the first column the conditions for comparison are represented, the symbols for which have been previously explained. The symbols, *E*, *S*, and *L*, in the second column, signify the judgments, equal to, shorter than, and longer than, respectively. In the next three columns are shown the percentages of judgments for the three durations used, namely, 5 seconds, 10 seconds, and 15 seconds. The percentage of judgments for the totals of the three durations are presented in the last column.

An inspection of the tabular data indicates that, for the 10- and 15-second durations, the results are consistent for the *NR* and *TR* comparisons. When *N* is taken as the standard, *R* is judged to be shorter, and when *R* is used as the standard, *N* is judged to be longer, which results are consistent. Likewise, when *T* is used as the standard, *R* is judged to be shorter, and when *R* is the standard, *T* is judged to be longer. The totals indicate the same tendency, when the above comparisons are made. When the *N* and *T* conditions are compared, it is observed that, for the 10- and 15-second durations, when *N* is taken as the standard, *T* is judged to be longer, and likewise *N* is judged to be longer when *T* is taken as the standard. Thus it is doubtful if the *T* condition is effective in influencing the judgments of the subjects under these experimental conditions. For the

5-second duration, the judgments fluctuate, and it is difficult to make any definite statements concerning the influence of either the *R* or the *T* condition.

The reliability of the data, which is to be found in Table 4*a*, was determined by the following formula to be found in Yule: $E_{12}{}^2 = p_1 \frac{q_1}{n_1} + p_2 \frac{q_2}{n_2}$. The formula is designed to show the reliability of the difference between an observed proportion and a theoretical proportion. $p_1\ q_1$ refer to the theoretical proportions, and $p_2\ q_2$ refer to the observed proportions. In the table, the first two columns indicate the various conditions for comparison. The third column shows the difference between the observed proportion and the theoretical proportion. For example, .33 is taken as the theoretical proportion, whereas .35 is the observed proportion. The difference between the two is .02. The $\sigma_{diff.}$ is indicated in the fourth column and the ratio of the difference between the observed proportion and the theoretical proportion to the sigma of the difference is shown in the fifth column. The ratios for the three durations and the totals are presented in the table.

If we take as the criterion of reliability a difference which is three times its σ, it will be observed that few of the differences are reliable. The re-

TABLE 4

PERCENTAGE OF JUDGMENTS FOR EACH CONDITION

Comparisons	Judgment	Duration in Seconds 5	10	15	Total
NR	*E*	.35	.25	.24	.28
	S	.47	.44	.49	.47
	L	.18	.31	.25	.25
RN	*E*	.42	.33	.22	.32
	S	.39	.15	.24	.26
	L	.19	.52	.54	.42
TR	*E*	.22	.17	.24	.21
	S	.64	.59	.51	.58
	L	.14	.24	.25	.21
RT	*E*	.53	.33	.21	.35
	S	.26	.17	.19	.21
	L	.21	.50	.60	.44
NT	*E*	.46	.29	.30	.35
	S	.28	.29	.26	.28
	L	.26	.42	.44	.37
TN	*E*	.26	.32	.39	.32
	S	.59	.26	.17	.34
	L	.15	.42	.44	.34

liability is influenced materially by the number of cases, thus the totals show greater reliability.

The same formula, indicated above, may be employed to determine the reliability of the difference between two observed proportions, as well as between an observed proportion and a theoretical proportion. For example, referring to Table 4, we may be concerned with the reliability of the dif-

TABLE 4*a*

RELIABILITY OF DATA IN TABLE 4

Duration	Condition	Judgment	Observed proportion minus theoretical	$\sigma_{diff.}$	Ratio
5 Seconds	*NR*	*E*	.02	.078	.25
		S	.14	.08	1.75
		L	.15	.07	2.14
	RN	*E*	.09	.079	.11
		S	.06	.079	.11
		L	.14	.08	1.75
	TR	*E*	.11	.073	1.50
		S	.31	.079	3.92
		L	.19	.063	3.02
	RT	*E*	.20	.08	2.50
		S	.07	.075	.93
		L	.12	.073	1.64
	NT	*E*	.13	.08	1.62
		S	.05	.076	.66
		L	.07	.075	.93
	TN	*E*	.07	.075	.93
		S	.26	.079	3.29
		L	.18	.076	2.37
10 Seconds	*NR*	*E*	.08	.075	1.06
		S	.11	.08	1.37
		L	.02	.077	.26
	RN	*E*	.00	.078	.00
		S	.18	.077	2.35
		L	.19	.08	2.37
	TR	*E*	.16	.07	2.28
		S	.26	.079	3.29
		L	.09	.081	1.11
	RT	*E*	.00	.078	.00
		S	.16	.07	2.29
		L	.17	.08	2.12
	NT	*E*	.04	.076	.52
		S	.04	.076	.52
		L	.09	.08	1.12
	TN	*E*	.01	.078	.13
		S	.07	.075	.92
		L	.09	.08	1.12

TABLE 4*a* (*cont.*)

Duration	Condition	Judgment	Observed proportion minus theoretical	$\sigma_{diff.}$	Ratio
15 Seconds	*NR*	*E*	.05	.076	.66
		S	.16	.08	2.00
		L	.06	.075	.80
	RN	*E*	.11	.073	1.51
		S	.09	.074	1.21
		L	.21	.08	2.62
	TR	*E*	.09	.074	1.21
		S	.18	.08	2.25
		L	.08	.075	1.06
	RT	*E*	.12	.07	1.71
		S	.14	.07	2.00
		L	.27	.079	3.42
	NT	*E*	.03	.077	.39
		S	.07	.075	.93
		L	.11	.08	1.37
	TN	*E*	.06	.08	.75
		S	.16	.07	2.29
		L	.11	.08	1.37
Total	*NR*	*E*	.05	.044	1.14
		S	.14	.046	3.04
		L	.08	.042	1.90
	RN	*E*	.01	.044	.23
		S	.07	.042	1.66
		L	.09	.042	2.14
	TR	*E*	.12	.041	2.92
		S	.25	.046	5.46
		L	.12	.041	2.92
	RT	*E*	.02	.044	.45
		S	.12	.033	3.63
		L	.11	.048	2.29
	NT	*E*	.02	.044	.45
		S	.05	.044	1.14
		L	.04	.044	.91
	TN	*E*	.01	.044	.23
		S	.01	.045	.22
		L	.01	.045	.22

ference between the *E* and *S* judgments, the *E* and *L* judgments, and the *S* and *L* judgments. In other words, using the figures in the total column as an example, we may ask: Is the difference between the two observed proportions, .28 and .47, a reliable difference or a difference due merely to chance? The reliability of the totals has been computed and the data are shown in Table 4*b*. For the *NR*, *RN*, *TR*, and *RT* comparisons, 7 of the 12 comparisons are reliable in accordance with the criterion stated in the above paragraph.

TABLE 4*b*

RELIABILITY OF DATA IN TABLE 4

Comparisons		Diff.	$\sigma_{diff.}$	Ratio
	E-S	.19	.053	3.6
NR	*E-L*	.03	.042	.71
	S-L	.22	.045	4.89
	E-S	.06	.043	1.39
RN	*E-L*	.10	.046	2.17
	S-L	.16	.045	3.55
	E-S	.37	.043	8.60
TR	*E-L*	.00	.036	.00
	S-L	.37	.043	8.60
	E-S	.14	.043	3.26
RT	*E-L*	.09	.047	1.91
	S-L	.23	.044	5.23
	E-S	.07	.044	1.59
NT	*E-L*	.02	.046	.43
	S-L	.09	.045	2.00
	E-S	.02	.045	.44
TN	*E-L*	.02	.045	.44
	S-L	.00	.046	.00

SUMMARY AND CONCLUSIONS

The experiments were preliminary in character, and we were concerned primarily in developing a technique of experimentation. The following tentative conclusions, however, seem to be justified on the basis of the objective data:

1) When the subjects were asked to estimate in terms of seconds the time that had elapsed, we find that the *R* condition (duration filled with music) is consistently judged to be shorter than the *N* and *T* conditions. The differences between the *N* and *T* conditions are relatively small for the 30-second and 45-second durations. We may assume that the judgments of the subjects were influenced by a rhythmical factor; however, whether it is a rhythmical condition that influences the estimation of time or the character and amount of filling cannot be definitely determined from the present experiment.

2) When the method of paired comparisons was used, the same tendency, i.e., for the subjects to judge the *R* condition as shorter than the *N* and *T* conditions, seems to prevail for the 10- and 15-second durations, and also for the totals. No definite statements can be made relative to the *NT* and *TN* comparisons. In general, these results appear to be consistent with those of the first experiment.

3) There is some evidence that the effectiveness of the *R* condition may be a function of the duration used, since the results of the second experiment for the 5-second duration are less consistent.

Colgate University
Hamilton, New York

THE ENERGY CONCEPT IN MODERN PSYCHOLOGY

David W. Oates

The place of the energy concept in psychology is a question of increasing importance: the fact that the conception of nervous and mental energy was chosen as the subject for a symposium at the International Congress of Psychology in 1923 is an indication of this growing interest. The necessity for introducing any hypothesis of energy to explain mental phenomena is disputed by some psychologists and as definitely supported by others. Adrian, for example, at the International Congress of Psychology, declared that, if the term "energy" was to be used in its purely physical sense, the conception of nervous energy was unnecessary: to this denial of its validity and necessity he added, however, "I am quite ready to believe that the conception of mental energy, properly defined, may be as necessary to psychology as that of physical energy is to physiology" (1, p. 125). Myers, on the other hand, fully accepts the energy concept, and writes, "I see no reason why we should not identify central nervous energy with mental energy" (19, p. 152). The arguments of the rival schools of thought on the matter need not be considered here, as Spearman (22) has dealt in some detail with the nature and history of the hypothesis in an admirable chapter in his recently published valuable volume on *The Abilities of Man.* Nor need we here discuss the mechanistic and vitalistic theories of energy production. The former would explain mental phenomena in terms of neuron activity, while the latter would regard consciousness and intelligence as independent data upon an entirely different plane from the physical. As Marston (12) has well put it, to the mechanistic set it is "the study of the influence of simpler energy units upon more complex energy units." The vitalist, on the other hand, holds "that higher energy units are the causes and that simpler energy units are effects" (pp. 18-19).

For our purposes here we might accept the position as stated by Thurstone (24) "In the last analysis the datum for psychology is the dynamic living self and the energy groups into which it may be divided . . . . These energy groups would be our innate, dynamic and more or less distinct sources of conduct, and we might come to call them drives, motives, instincts, determining tendencies, or any other word which represents that which we as individuals innately really are" (p. 11). The generally accepted physiological explanation of nervous energy is that given by McDougall (13; 14,

p. 170; 15, p. 161) who regards the constituent neurons of the nervous system "as a vast system of channels in all parts of which potential chemical energy is constantly being transformed, in virtue of the normal vital activity of the neurones, into a particular form of active energy. This energy . . . . tends always to flow . . . . from places of higher to places of lower potential, following the paths of least resistance." But, as Marston (12) has indicated, there is in consciousness a stabilized form of energy over and above the nervous impulse. Psychologists, neurologists, psychophysiologists and possibly psychoanalysts, he maintains, substantially support this theory. "All these types of investigators assume, either tacitly or explicitly, that consciousness is a manifestation of energy which exists and reacts as a unit separate from mere intra-neuronic disturbance" (p. 26).

A theory of psychic energy has been developed by Jung in his "libido" which he regards as the equivalent of physical energy in the psychical sphere. Ray (20) in a paper on "Life Energy," in which he speaks of the living body as an "energy transformer" and the brain and spinal cord as "transformers of that particular kind of energy we group under the term 'mental'" extends the energy concept. Gillespie (9) has examined some of the forms the energy concept has assumed in the hands of certain other psychologists. It is the purpose of this short note to indicate the essential place of the concept of energy in four recent psychological theories as an illustration of the value of the hypothesis of psychic energy as a starting point for the elaboration of these theories and to suggest a possible basis for the fundamental similarity of form assumed by the concept in these theories. The theories are so well established in psychological thought that a very brief statement of them will be sufficient for our purpose and we may then examine the similarity of emphasis upon the hypothesis which they present.

1) *Spearman's Two-Factor Theory of Intelligence* (22, 23). Spearman maintains that there is a central factor, *g,* in all our intellectual abilities which may be roughly identified with "general intelligence," and, side by side with this general factor, are numerous specific factors. He regards his general factor as "a central fund of intellective energy," and explains it by postulating an energy which flows from the whole cortex. His specific factors are explained on the analogy of engines as the various mechanisms which this energy drives.

2) *Downey's Will-Temperament.* The temperamental make-up of an individual, according to this theory, is determined by two factors, (*a*) the amount of nervous energy at the disposal of the individual, and (*b*) the tendency of such nervous energy to discharge immediately into the motor areas that innervate the muscles and glands; or, on the contrary, to find a way out by a roundabout pathway of discharge (7, p. 59).

3) *McDougall's Theory of Instincts.* This can be stated in his own words. "The essence of instinctive activity seems to be such liberation and

direction of energy which perhaps we may best speak of as psycho-physical energy . . . . we might regard the several instincts as somehow drawing upon a common store of reserve energy . . . . We might regard each instinct as containing a store of potential energy . . . . Some combination of these two views seems to be required" (16, pp. 107-108). Here we have the hypothesis of a common fund of energy upon which all of the instincts can draw, and a specific store of energy attached to each instinct.

4) *Burt's Theory of General Emotionality* (4; 5, pp. 506-507). The emotional nature of an individual according to this theory appears to be due to, first, a central factor which may be termed "General Emotionality," and is identified with "the total energy at the disposal of the individual" (6) and, secondly, to a number of more specific emotional reactions. That is, emotional qualities may be analyzed into a broad single central factor of general emotionality and a number of specialized tendencies based on specific instincts or emotions.

When these theories are stripped of conflicting terms it would appear that they have certain features in common, and this similarity, it is suggested, has sufficient significance to be worthy of consideration. The four theories postulate a general fund of energy. Spearman definitely terms it a central fund of energy. Downey's differentiating factor between individuals is the amount of nervous energy, again a central fund of energy. McDougall assumes a general fund of energy upon which all the instincts may draw; and Burt's general emotionality depends upon a general fund of emotional energy. The parallelism goes still further: it is assumed that this energy, the amount of which available in any individual is more or less fixed, is capable of flux and of change of direction or freedom of flow. The specific factor is implied in the intellectual field by Spearman's hypothesis of "engines," which give qualitative differentiation to the quantitative general factor by modifying the direction or nature of the flow of the energy. In Downey's theory certain inhibiting factors modify the flow of the sum total of psychic energy available. The modifying factors operating upon the central fund of energy in the theories of McDougall and Burt may also be related to the concept of psychic energy in very similar ways: in the former case it will depend upon the working of the specific instincts which vary in the relative degree of strength in the energy attached to them; and, in the second case, in a like manner a specific emotion may cause or relax a tension, thus blockng or clearing the path of discharge of the flowing energy. It appears that at least in the last three theories the modifying factor in its functioning is related to McDougall's theory of the vicarious usage of freed nervous energy, and that modification arises through the drainage of energy from one part of the nervous system to another. And even in the case of the first theory the efficiency of Spearman's engines will be determined by the relative ease and smoothness with which they allow the central intellective energy to flow and thus express itself.

We have, therefore, in the theories under examination a common hypothesis of a reservoir of freed energy and a modifying factor determining the direction and rate of flow of the psychic energy available.

The parallelism that, as we have briefly indicated, may be traced in these theories appears to point to a common nucleus. The psychological entities in which this common nucleus can find a basis are, it is suggested, the instincts which are regarded by many psychologists today as the driving forces of all human activity. This contention will apparently conflict with the views of a group of social psychologists, chief of whom are Allport, Dunlap, and Bernard, who have rejected the concept of instinct. The position of the instinct-deniers has been clearly stated and their view analyzed by Kuo. Woodworth also has critically examined the arguments for the elimination of the concept and shown that they are largely irrelevant. It is not the purpose of this paper to enter into the instinct controversy, the outcome of which will not materially affect the validity of our conclusion. Dunlap (8, p. 16) replaces instincts by desires. Allport substitutes prepotent reflexes and prepotent habit "which acquires a compelling power similar to that of the prepotent reflexes" (2, p. 109). Bernard substitutes drives which result from internal organization and adds, "the drive is basically an organic fact and bears a close relationship to what Allport calls the prepotent impulses" (3, p. 213). Wyatt sums up an examination of the arguments with a remark that if objection be made "that there still remains a wide difference between these drives, desires and impulses, and the instincts of the older psychology, in that the latter are more complicated or variable in their reaction-patterns, and assume an act of perception, while the drives and impulses are simpler and do not involve consciousness, the answer is that the distinction vanishes on examination" (26, p. 131). However that may be, it will be generally agreed that the desires, impulses, and drives are sources of the energy involved in any reaction and these may, therefore, be classed with the instincts as constituting the psychological basis of the energy which forms the common element in the hypotheses here under examination.

The explanation we offer is least obvious in the case of the theory of intelligence. The main facts, however, seem to indicate that intelligent activities have developed from activities based on the instincts, and, as Ernest Jones has remarked "one can see no dynamic significance whatever in ideas except in so far as they function as the representative of some impulse or other" (10, p. 259). The explanation here suggested is in accord with McDougall's views. "Directly or indirectly," he writes, "the instincts are the prime movers of all human activity; by the conative or impulsive force of some instinct (or some habit derived from an instinct) every train of thought is borne along to its end. . . . These impulses are the mental forces that maintain and shape all the life of individuals, and in them we are confronted with the central mystery of life and mind and will" (17, 13th ed., p. 44). Mukherji (18) has also shown that psychic and neural energies are two differentiated forms of biological energy, and that this libido or inborn neuro-psychic energy supplies the functional energy to all the instincts.

The all-embracing claims of McDougall's last phrase have been challenged; Reid (21), for instance, questions whether the moral life can be wholly explainable in terms of instinct, or whether at the human level there is not something omitted from the theory of instinct psychology. We are not here concerned to examine the relationship of instinct to will and the part played by instinct in moral development, we are only interested to note the relationship of the four theories under consideration, which are based upon essentially similar forms of the energy concept, to the central fact of the instincts which, as the source of this postulated fund of energy, appear to supply the basis for the common nucleus which we have traced in these theories.

REFERENCES

1. Adrian, E. D. The conception of nervous and mental energy. *Brit. J. Psychol.,* 1923, **14**, 121-125.
2. Allport, F. H. Social psychology. Boston: Houghton Mifflin, 1924. Pp. xiv+453.
3. Bernard, L. L. Introduction to social psychology. New York: Holt, 1926. Pp. x+651.
4. Burt, C. General and specific factors underlying the primary emotions. *Rep. Brit. Asso. Adv. Sci.,* 1915.
5. ———. The young delinquent. London and New York: Appleton, 1925. Pp. 619.
6. ———. (Letter.) *Brit. J. Psychol.,* 1926, **16**, 215.
7. Downey, J. E. The will-temperament and its testing. Yonkers, N. Y.: World Book Co., 1924. Pp. vi+339.
8. Dunlap, K. Social psychology. Baltimore, Md.: Williams & Wilkins, 1925. Pp. 261.
9. Gillespie, R. D. Concepts of nervous and mental energy. *Brit. J. Psychol.,* 1925, **15**, 266-279.
10. Jones, E. The classification of instincts. *Brit. J. Psychol.,* 1924, **14**, 256-261.
11. Kuo, Z. Y. The net result of the anti-heredity movement in psychology. *Psychol. Rev.,* 1929, **36**, 181-199.
12. Marston, W. M. Materialism, vitalism and psychology. *Psyche,* 1928, No. 31, 18-34.
13. McDougall, W. Physiological factors in the attention process. *Mind,* 1902, **11**, 316-351; 1903, **12**.
14. ———. Inhibitory processes within the nervous system. *Brain,* 1903, **26**, 153-191.
15. ———. Physiological psychology. New York: Macmillan, 1905. Pp. xiii+172.
16. ———. An outline of psychology. New York: Scribner's, 1923. Pp. xvi+456.

17. ———. An introduction to social psychology. London: Methuen, 1908; Boston: Luce, 1909.
18. Mukherji, K. C. The biological concept of libido. *Amer. J. Psychol.,* 1926, **27**, 553-556.
19. Myers, C. S. The conception of nervous and mental energy. *Brit. J. Psychol.,* 1923, **14**, 148-152.
20. Ray, M. Life energy. *Psyche,* 1928, No. 31, 51-59.
21. Reid, L. A. Instinct, emotion and the higher life. *Brit. J. Psychol.,* 1923, **14**, 76-93.
22. Spearman, C. E. The abilities of man. London: Macmillan, 1927.
23. Spearman, C. E., & Hart, B. General ability; its existence and nature. *Brit. J. Psychol.,* 1912, **5**, 51-84.
24. Thurstone, L. L. The nature of intelligence. New York: Harcourt, Brace; London: Kegan Paul, 1924. Pp. 167.
25. Woodworth, R. S. A justification of the concept of instinct. *J. Abn. & Soc. Psychol.,* 1927, **22**, 3-11.
26. Wyatt, H. G. The anti-instinctive attitude. *Psychol. Rev.,* 1927, **34**, 126-132.

Municipal Secondary School for Boys
Newport, England

APPARATUS

A DESCRIPTION OF THE LECTOMETER AND THE RELIABILITY OF LECTOMETER SCORES*

WILLARD L. VALENTINE AND MAX F. MEYER

The lectometer (Latin, *lectum,* read) is one of a group of instruments, exhibiting scores of low intercorrelation, first used in research in vocational guidance at the University of Missouri. The first model was demonstrated by Meyer at the Wisconsin meeting of the American Psychological Association, December, 1923. A second model was built in the Ohio State Laboratory during 1928. It is the construction of this model which is described in this paper. All of the essential features of the original were retained in the copy so that the scores on the two instruments are comparable.

The task imposed by the instrument forces the subject to copy the dots from the row of blocks onto the row of tablets (see Figure 1). Copying is done by pushing, one at a time, the switchboard switches composing the tablets. This is about as simple a kinaesthetic task as can be arranged, so that the critical, primary source of sensory guidance is visual. Not visual in the sense that refractive errors or poor acuity would make for poor scores, but rather in the sense that the reactions of some individuals to visual stimuli are more numerous and more detailed than his reactions to non-visual stimuli. The separation of that which Meyer has called the *theatheratic* (Greek: *thea-theran, sight-hunting*) component in human behavior from the structural deficiencies and abnormalities of the visual apparatus has been neglected in recent years and seems to have been recognized historically only with respect to individual differences in imagery types.

In order to standardize somewhat the amount of equivalent previous practice in the task, the construction of the apparatus requires that the subject copy the dots onto the tablets in accordance with the verbal instructions: "For every dot in the *top* row on the block, push a key in the *bottom* row in the tablet and for every dot in the *bottom* row on the block push a key in the *top* row in the tablet." This assures that the subjects will have had a minimum amount of usable previous practice. Without these instructions certain vocational groups would enter the test with an advantage due to their vocational habits. As it is, only printers have the advantage.

---

*This study is a part of a program of research on the psychology of the highway to which grants have been made by the Division of Anthropology and Psychology of the National Research Council from funds provided by the American Optimetric Association.

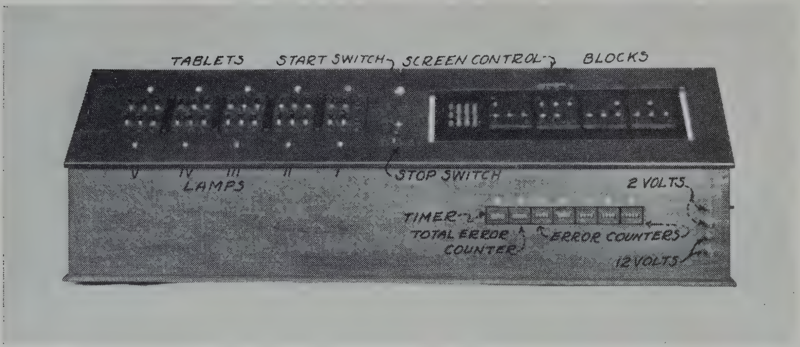

TABLETS
START SWITCH
SCREEN CONTROL
BLOCKS
V
IV
III
II
I
LAMPS
STOP SWITCH
2 VOLTS
TIMER
TOTAL ERROR COUNTER
ERROR COUNTERS
12 VOLTS

FIGURE 1

The Lectometer from the Rear

The screen removed showing the first four blocks in place. Over-all dimensions, 9″ x 10″ x 48″.

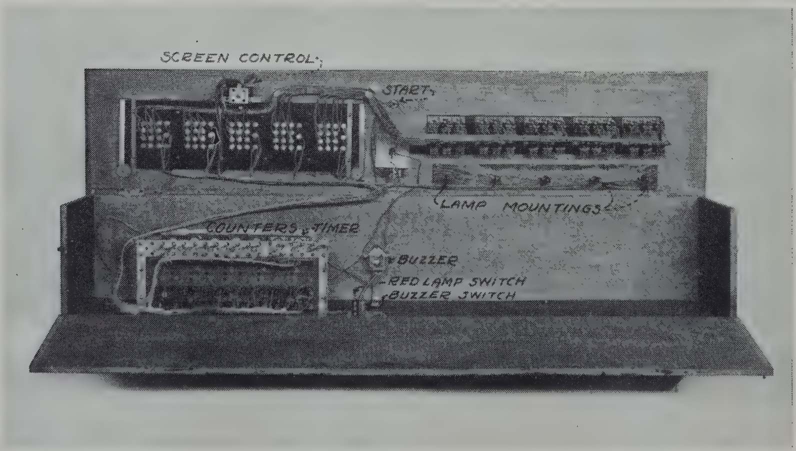

SCREEN CONTROL
START
LAMP MOUNTINGS
COUNTERS
TIMER
BUZZER
RED LAMP SWITCH
BUZZER SWITCH

FIGURE 2

The Lectometer from the Front

The top raised and the front lowered showing the wiring details. A tray of extra blocks fits in the empty space to the right.

The symbols used are characters in the American Point system for the blind. All are some arrangement or variation of six dots in two horizontal rows, as $\overset{\cdot\;\cdot\;\cdot}{\cdot\;\cdot\;\cdot}$. The letter V is $\overset{\cdot\quad\cdot}{\cdot\quad}$ and would be copied $\overset{\quad\cdot}{\cdot\quad\cdot}$. The lamp (Figure 1) will light up only when the letter is completed, and if an error is made a buzzer sounds, or a red lamp lights up. The subject is instructed to correct all errors as soon as they are made. The blocks must be copied onto the corresponding tablets from left to right in order. The time required to light the five lamps and to correct any errors that have been made is the score.

A Description of the Wiring Diagram

The timing mechanism may be either a clock or a pendulum arranged to make contact in any convenient time unit.[1] We have used 1-, ½-, and 1/5-second units. The 1-second unit is probably precise enough. The 2-volt battery circuit (Figure 3) actuates the vibrator relay which makes contact 2. Thus the timer receives a 12-volt pulse each time unit as long as contact 3 is closed. Considering the major 12-volt battery circuit, the stop-key is closed by the experimenter at the beginning of a test. The subject presses the push-button marked "Start" at his convenience. This does two things: It releases the screen so that it can fly back, by means of a spring mechanism not shown, and at the same time makes contact 5, which starts the time recorder. Contact 5 remains closed by reason of the self-locking feature of the relay. As soon as 5 is made, the first bank of mercury cups is supplied with current by way of 10 and cup 7. The blocks make contact with the cups by means of wire nails (brass escutcheon pins) projecting from the bottom, one for each cup. The 12 cups not numbered are to be considered terminals of the six possible dots. Those pins that are to form the pattern are wired in series inside the block so that when the corresponding switches are closed, a circuit is made through mercury cup 11. Those that are to be omitted from the pattern are wired in parallel with 7 and 8, so that when *any one* switch corresponding to them is closed, a circuit is completed through 8 which trips counter I and the error relay. When the error relay is energized, contact 9 is closed and the buzzer and total error counter are actuated. Thus, whenever an error is made, two counters operate; counter I only for errors made in the first block and the "total error counter" for any error in the five blocks. If one is not interested in the distribution of errors over the separate blocks, the five individual counters may be eliminated.

It will also be noted that the circuit from 11 to lamp I will be completed only as long as the break contact (17), attached to the armature of the

---

[1] Where it is possible to obtain maintained, 60-cycle current, the previously described timing mechanism may be used [see (2)].

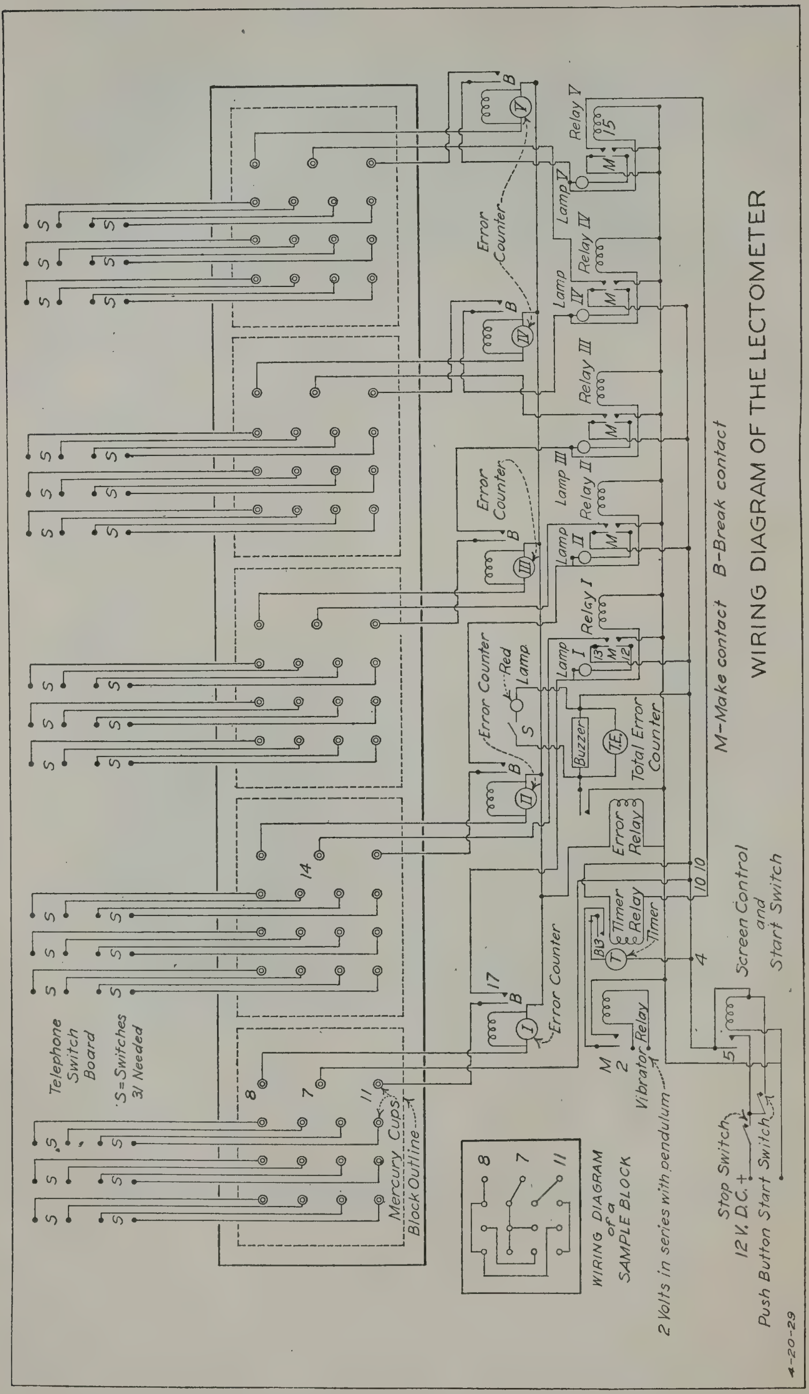

Telephone Switch Board
S=Switches 31 Needed
S
Mercury Cups
Block Outline
8
7
11
14
17
B
Error Counter
Red Lamp
S
I
II
III
IV
V
Lamp I
Lamp II
Lamp III
Lamp IV
Lamp V
Relay I
Relay II
Relay III
Relay IV
Relay V
M
12
13
15
Buzzer
T.E.
Total Error Counter
Error Relay
Timer Relay
Timer
T
B13
10 10
4
5
Screen Control and Start Switch
Relay
M
2
Vibrator
2 Volts in series with pendulum
WIRING DIAGRAM of a SAMPLE BLOCK
8
7
11
Stop Switch
12 V. D.C. +
Push Button Start Switch
4-20-29
M-Make contact B-Break contact
WIRING DIAGRAM OF THE LECTOMETER

FIGURE 3

The Wiring Diagram of the Lectometer

counter, is closed. As any error in copying block I will pull the armature up, then lamp I will not light until all errors in block I have been corrected.

From cup 11, relay I makes two contacts: 12, which throws lamp I in parallel with the relay, and 13, which supplies cup 14, throwing the second bank of cups in parallel with the first. Thus it is required that the blocks be copied in order; block II is "dead" until I has been copied. From here the process is repeated for as many blocks as desired. The last relay makes contact 15, which closes the circuit through the timer relay and breaks contact 3, thereby stopping the timer. As each lamp will light up only after all errors in copying on the corresponding tablet have been corrected, the timer will stop only after all errors made any place in the system have been corrected.

After the subject finishes the problem, the experimenter opens the stop switch putting out all the lamps and permitting the screen to be locked in place over the blocks. He makes readings on the battery of counters and pushes back the switches in the tablets.

Accessory Details

The counters have been previously described by Miles (1), but the relays are not in common use in psychological laboratories. They are manufactured for telephone exchange installations and may be procured through any telephone supply company. They are much smaller than the pony type and may be mounted in banks in a small space. Each coil actuates three make and three break platinum contacts as delivered, but they can be made to operate 12, 18, or 21 contacts with little trouble. Each two coils is mounted as a unit with its contacts inside a dustproof cover. After a minimum of adjustment they operate for months without attention. The magnets are highly efficient and may be had in resistances ranging from 100-2500 ohms. If desired, the relays can be furnished with magnets fitted with a copper head in which an induced current will hold the armature for an appreciable time after contact through the coil is broken.

The switches have not been previously described. They are used commercially in telephone switchboards. They were adopted after a year's experiment with the General Electric switch 1299. The latter switch will not stand up under the hard usage of an extended investigation. The switchboard switches possess further advantages which make them adaptable in psychological instruments. They can be adjusted to perform either single- or double-throw functions and will make either momentary or permanent contact on either side as double throw. As delivered they make two contacts and break two on both sides. More makes and breaks can be provided as the problem requires.

The Testing Procedure

A group of five blocks is placed under the screen (Figure 1). The subject is instructed according to the verbal formula (cf. above) and permitted to perform a preliminary test. This time is not counted in the score. If he makes no errors, he is instructed to purposely make some and correct them.

A new set of blocks is placed on the instrument without the subject's having an opportunity to observe them. He is then instructed to work as quickly and accurately as possible. This test is done twice, the sum of the scores for the two times becoming his score for Test I. The two parts of Test I, called *a* and *b*, are recorded separately. In the same way he completes Tests II, III, IV, V, and VI, the blocks being changed between each test. The worst score is cast out and the sum of the remaining five tests (10 performances) becomes the score used in the distributions.

In a series of auxiliary experiments, it was determined that there was no significant difference in the difficulty of the blocks as measured by the number of errors made on each. The position of the block in the series is a much more severe determinant of the number of errors that will be made. Those coming first exhibit the most errors. The worst score (the one cast out) was the score on Test I in 30% of 156 unselected cases, Test II in 40%, Test III in 17%, Test IV in 6%, Test V in 3%, and Test VI in 4%. In order to keep the length of the tests constant, it is necessary to keep the total number of dots on all five blocks approximately constant. The number 15 best serves this end and was consequently adopted as a standard. After this preliminary work, the five blocks decided upon for each test were fastened together as a unit and changed in one operation, materially reducing the time required to give the test.

The Experimental Group

The college freshmen and sophomores who served as subjects fall in three classifications: Group I consisted of those who took the lectometer test once after a series of learning experiments on the instrument. The 34 in the group were thoroughly acquainted with the apparatus and technique. Group II consisted of 54 men and 54 women who took the test once, never having seen the instrument prior to that time. Their data, besides furnishing norms, were used to compute a reliability coefficient between the two halves of the test. Group III consisted of 34 men and 28 women. These took the test twice, at an interval of about a month, furnishing data for another computation of the reliability coefficient. Patently, there is no difference between Groups II and III as far as the first test for the latter is concerned. In order to obtain more reliable coefficients, then, they may be combined into Group II*a* consisting of 88 men and 82 women.

Results

The reliability of the test, after extended practice on the instrument, is expressed by a Personian coefficient of correlation between the two halves of the test of .727±.09 for 13 men and .850±.04 for 24 women. These are to be considered in the light of the small number of cases. The most

TABLE 1

THE STATISTICAL CHARACTERISTICS OF THE FOUR GROUPS

Group	Men	Women
Group I		
N	13	21
$r_{½½}$	0.727±0.09	0.850±0.04
Spearman-Brown r	0.842	0.919
Mean score	107.8	120.8
$S.D._{dist.}$	23.0	31.6
$S.D._m$	6.4	6.9
Group II		
N	54	54
$r_{½½}$	0.924±0.01	0.888±0.02
Spearman-Brown r	0.961	0.941
Mean score	136.3	152.4
$S.D._{dist.}$	37.2	49.1
$S.D._m$	5.1	6.7
Group II*a*		
N	88	82
$r_{½½}$	0.913±0.01	0.904±0.01
Spearman-Brown r	0.954	0.949
Mean score	134.3	157.5
$S.D._{dist.}$	34.9	60.4
$S.D._m$	3.7	6.7
Group III		
N	34	28
Test A: $r_{½½}$	0.899±0.02	0.926±0.02
Test B: $r_{½½}$	0.816±0.04	0.951±0.01
r_{AB}	0.784±0.05	0.826±0.04
Test A: Spearman-Brown $r_{½½}$	0.947	0.962
Test B: Spearman-Brown $r_{½½}$	0.898	0.975
Test A: mean score	131.1	166.0
Test B: mean score	100.9	123.3
Test A: $S.D._{dist.}$	34.9	90.7
Test B: $S.D._{dist.}$	18.6	40.5
Test A: $S.D._m$	6.0	17.2
Test B: $S.D._m$	3.2	7.7

significant item in the data of Group I is the mean score, which compares favorably in each case with the corresponding mean of Group III, Test B. The combined mean of the men and women in Group I is 115.8, and the combined mean of Group III is 111.1. This agreement indicates that verbal instruction plus one practice is approximately equivalent to a long series of practices without instruction. It also indicates that the mean limiting performance time is located in the neighborhood of 100 seconds. The shortest time recorded is 67 seconds.

For naïve subjects the reliability is .913±.1 for 88 men and .904±.01 for 82 women when the r is computed between the two halves of the test. As expressed by r_{AB}, where A and B are two tests separated by an interval of one month, the coefficient is .784±.05 for 34 men and .826±.04 for 28 women.

TABLE 2

Sex Differences in the Four Groups

Group	Sex	No.	Mean score	*Diff.*	$P.E._{diff.}$	$\frac{Diff.}{P.E._{diff.}}$
I						
	Men	13	107.8			
	Women	21	120.8	13.0	6.3	2.1
II						
	Men	54	136.3			
	Women	54	152.4	16.1	5.6	2.9
II*a*						
	Men	88	134.3			
	Women	82	157.6	23.3	5.2	4.5
III						
	Men	34	131.1			
	Women	28	166.0	34.9	12.3	2.9

Sex differences are consistent in all indices with respect to amount and direction. Table 2 summarizes these data. The largest group gives a reliable difference expressed by 100 chances in 100 that the difference is a result of some factor, or group of factors, in the experiment and not due to chance. Other sex differences between the means of each half of the test are of the same order of significance, none falling below 92 chances in 100 that the difference would be substantiated upon repetition.

It is extremely likely that these sex differences could be eliminated by the proper experimental procedure. One could, for example, give the test to pairs of subjects, instructing them that the purpose of the experiment was to find out whether or not men were better than women.[2]

---

[2] It is beyond doubt, however, that without this technique it is necessary to record the scores of men and women separately. For instance, the 50th centile of the men (124 sec.) is the 72nd centile of the women.

SUMMARY

1) The lectometer has a reliability expressed by several Pearsonian coefficients ranging from .727±.09 to .926±.02. The most convincing of these, from the point of view of the number of cases, is an $r_{\frac{1}{2}\frac{1}{2}}$ of .913±.01, and, from the point of view of technique, an r_{AB} of .784±.05.

2) The mean score of the women indicates poorer performance.

REFERENCES

1 MILES, W. R. Note on electric counters. *J. Exper. Psychol.*, 1922, **5**, 76-77.

2. VALENTINE, W. L. A small, self-starting synchronous motor. *Amer. J. Psychol.*, 1928, **40**, 495-496.

Ohio State University
Columbus, Ohio

University of Missouri
Columbia, Missouri

BELL LABORATORY VICTOR RECORDS

PAUL R. FARNSWORTH

The Bell Telephone Laboratories have recently issued a set of four Victor records which should be of great interest to psychologists. This little battery appears to the writer as splendid for class demonstration. Music and speech are presented normally, and are then distorted in various ways.

Record No. B T L 1 displays the effects of screening out successively all frequencies below 375 d.v., below 1250 d.v., above 2500 d.v., and above 1250 d.v. The normal range is offered for the sake of contrast. B T L 2A shows the result of overloading the amplifier. Later in the record the amplifier is overloaded and frequencies below 375 d.v. and above 2500 d.v. are screened out. Face B of the record offers music and speech distorted by the elimination, first, of all frequencies below 375 d.v. and above 2500 d.v., and, later, of all below 750 d.v. and above 2500 d.v. Record B T L 3 will probably prove of least interest to psychologists, since it demonstrates the effects of sound repetition at different transmission unit levels. B T L 4A demonstrates quite clearly that the timbre differences of the piano, 'cello, and French horn depend upon overtones. Face B illustrates the interesting fact that a tone's fundamental may be eliminated without a change in its apparent pitch.

Mr. John Mills, Director of Publication for the Bell Telephone Laboratories, has informed me by letter that he hopes to prepare in the near future additional records which may be of interest and value to departments of psychology. He writes further in reference to the records I have already described: "This series of records we stand ready to send, with the compliments of these Laboratories, to departments of universities and colleges, but not to secondary schools, which would find them of value. The departments most likely to be concerned are, of course, physics, electrical engineering and psychology and occasionally music and public speaking.

"You may, therefore, count upon our willingness to send any of our phonograph records upon request to university or college departments of psychology. Where we do so we should very much like it to be understood that the records are in the nature of a loan which we will maintain in good working condition; that is, if any record is marred or scratched it should be returned for replacement, and if, from time to time, we produce better records an exchange is to be made."

Since this article went to press, three other records have been issued: B T L 5 demonstrates more fully the phenomena exhibited by B T L 4B; B T L 6 gives illustrations of phase distortion in speech, and B T L 7 indicates a method for measuring noise.

Stanford University
California

PHONOGRAPH RECORDS FROM THE BELL TELEPHONE LABORATORIES ILLUSTRATING PHENOMENA OF ACOUSTICS

ROBERT H. SEASHORE

A series of disk phonograph records is being produced by Bell Telephone Laboratories which illustrate some of the major phenomena of acoustics of interest to physicists and psychologists. The effects which are recorded require for their production so much complicated apparatus as ordinarily to be beyond the possibilities of academic laboratories of physics or psychology. A series of phonograph records was, therefore, initiated to make these effects available to educators for lecture and classroom work. The records, which are not for sale, have, however, been distributed with the compliments of the Laboratories, upon the request of college and university departments of physics, and it is understood that requests from similar departments of psychology would be met. Requests are most conveniently directed to the Bureau of Publication, Bell Telephone Laboratories, 436 West Street, New York, N. Y.

In each record the normal sound, a selection of instrumental music or speech, is compared with the same selection as modified by electrical instruments such as high-pass and low-pass filters for elimination of certain pitch frequencies, and underloaded and overloaded amplifiers and a phase changing device. The list of records so far produced and the experimental changes illustrated are as follows:

1A. Suppression of selected frequency bands from an orchestral selection to give
  - *a*) none below 375 v.d.
  - *b*) none below 1250 v.d.
  - *c*) none above 2500 v.d.
  - *d*) none above 1250 v.d.

1B. Same as 1A except for speech.

2A. Overloading of amplifiers and the same with elimination of frequency ranges below 375 v.d., and above 2500 v.d., using both music and speech.

2B. Limitation of bands of frequency to within 375 v.d. to 2500 v.d. and within 850 v.d. to 2500 v.d., both music and speech.

3A. Reduction of loudness of music by 1, 3, 5, 10, 20 and 30 transmission units.

3B. Same as 3A except for speech.

4A. Effect of eliminating higher frequencies from tones of piano, cello, and French horn, each playing middle C, 256 v.d. (This record is perhaps the one of greatest interest to psychologists as illustrating the analysis of timbre of different instruments by overtone relationships.) Illustrates successively
  - *a*) complete tone
  - *b*) fundamental only (1st partial)
  - *c*) first two partials
  - *d*) first three partials
  - *e*) first ten partials
  - *f*) first eighteen partials

4B. Same as 4A except effects of eliminating lower frequencies.

5A. Recognition of the pitch of complex vocal tones without fundamental and without lower partials up to the sixth.

5B. Same as 5A except for organ and clarinet tones.

6A. Phase distortions of speech component frequencies below and above 1000 v.d. by time differences of .10, .07, .035 seconds illustrating the importance of the short duration of speech sounds.

Two other records will shortly be sent out and descriptions of them is as follows:

7A and B. A description of the method of measuring noise by its masking effects on definite pitch ranges of audition; warble tones are illustrated, varying in intensity and pitch range; various noises are illustrated; what an observer would hear if a warble tone varied in intensity until just audible over the noise; and then the tone alone. (A poster illustrating the apparatus and typical audiograms is available upon request for wall display.)

8A and B. Acoustic reflections of speech from distant reflecting surfaces, first 35 feet and then 100 feet away, with and without attenuation. Reflections from both distances at once, with and without attenuation; two additional reflections; and similar reflections for music.

University of Oregon
Eugene, Oregon

CONTINUOUS QUANTITATIVE RECORDS IN STANDARD UNITS BY THE WHEATSTONE BRIDGE METHOD[1]

Chester W. Darrow

The *Behavior Research resistance box* is designed to be used auxiliary to a Wheatstone bridge for making continuous quantitative records of the galvanic skin reflex without repeated rebalancing or elaborate calculation. The advantages of the bridge method, extreme sensitivity, and measurement in standard units (ohms), are retained. The device may be used when registration is made photographically, when tracings are made by following the deflections by hand, or when record is kept by merely noting down the numerical magnitude of the changes.

Essentially the resistance box consists (*A*) of an adjustable high resistance which may be substituted for the subject in the *X* arm of the bridge; (*B*) of a switch for making this substitution; (*C*) of a variable standard resistance arranged in convenient (1000 ω) steps placed in series with the *X* resistance; (*D*) of an adjustable shunt for the galvanometer; and (*E*) of a reversing switch. In practice it has been found convenient to inclose these parts, with the knobs on the side, in a case beneath the regulation bridge. The parts are arranged schematically in panels around a standard bridge in Figure 1.

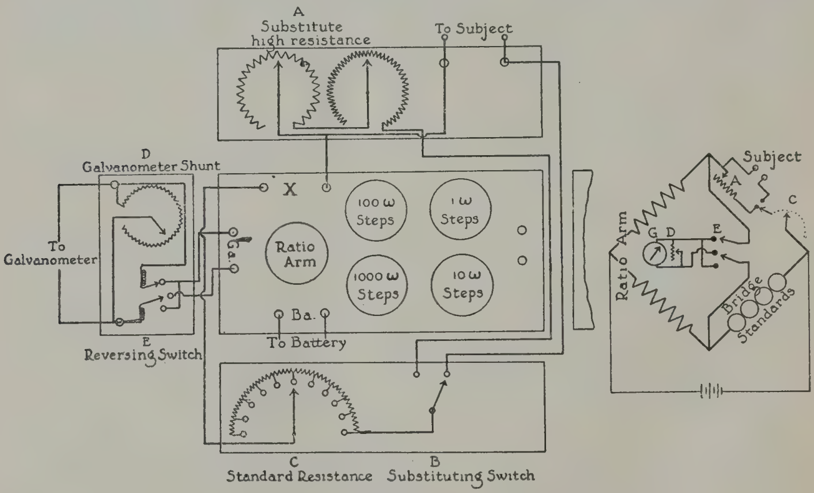

A
Substitute high resistance
To Subject
D
Galvanometer Shunt
To Galvanometer
X
Ga.
Ratio Arm
100 ω Steps
1 ω Steps
1000 ω Steps
10 ω Steps
Ba.
To Battery
E
Reversing Switch
C
Standard Resistance
B
Substituting Switch
Subject
A
C
Ratio Arm
G
D
E
Bridge Standards

FIGURE 1

The procedure with this device is as follows: (*a*) The subject is placed in the *X* arm of the bridge by switch *B*. (*b*) The variable standard resistance, *C*, is set to make a convenient (5000 ω) addition to the resistance of the subject. (*c*) The subject, plus the known standard, *C*, is balanced by adjustment of the bridge, and the reading and the position of the ratio arm noted. (*d*) The adjustable high resistance, *A*, is substituted for the subject by switch *B*, and is adjusted to balance the bridge, thus making this stable resistance and the original resistance of the subject equal. (*e*) The standard resistance, *C*, is reduced by a convenient amount (1000 ω); the shunt, *D*, is adjusted so that this change shows a convenient galvanometer deflection; the amount of the deflection is recorded; and then this resistance is returned to its original magnitude. We have now a record of the basic resistance of the subject and of the sensitivity of the galvanometer setting, and we may proceed, by a shift of switch *B*, but without changing the other settings, to obtain experimental records. When at any time the deflections in either direction become so great as to approach the range limit of the galvanometer, compensation is afforded by adding or subtracting standard units (1000 ω) from the resistance, *C*, in series with the subject, and noting the amount of these changes in the record. Such changes seldom exceed the five steps, plus or minus, of 1000 ω each, which have been provided for in our preliminary construction, and, therefore, no new balancing of the bridge is ordinarily required. If excessively strong excitation is to be employed, or if the subjects have very high initial resistances a standard with a wider range is desirable. We have recently added a series of 10,000 ω steps to the "C" resistance of our own equipment.

---

[1] Studies from the Behavior Research Fund and the Institute for Juvenile Research, Chicago. Series B, No. 159.

The following procedure[2] enables us to obtain in ohms the values of deflections obtained in the above fashion: (*a*) The bridge resistance and ratio arm are set at the respective values originally recorded. The ratio arm *must be in the original position.* (*b*) The variable high resistance, *A*, plus the known standard, *C*, is adjusted to balance the bridge. (*c*) The original change in the standard resistance, *C*, is duplicated, and the shunt, *D*, is adjusted so that this change gives the original deflection. Basic resistance and sensitivity are now as they were when the subject originally was in the circuit, and the variable high resistance, *A*, occupies the place of the subject in the *X* arm of the bridge. If this substitute resistance were calibrated[3] it would be possible to make adjustments duplicating all deflections of the subject and thus to obtain readings on all reactions. The same purpose is served by turning the reversing switch, *E*. The variable high resistance, *A,* now occupies the position relative to the galvanometer and the source of potential which was originally held by the variable resistance of the bridge, while the variable bridge resistance (to be multiplied by the value on the ratio arm) occupies the position of the subject. Adjustments of the bridge may now duplicate all galvanometric changes, and readings from the bridge may be made in ohms on all deflections. The standard resistance, *C*, in series with the *X* resistance must, of course, be subtracted. Where it is not convenient to run the record through the apparatus a second time, an empirical scale of the various unit deflections for any given setting of the apparatus is easily constructed and may be used as a measure.

---

[2] For assistance in the mathematical justification of this procedure the writer is indebted to Dr. Gulliksen of the Department of Psychology of the University of Chicago.

[3] The uncalibrated non-inductive variable high resistances by "Frost" which are used for the substitute resistance, *A*, and for the galvanometer shunt, *D*, were found compact and low in cost, and, with the arrangement here suggested, only one more expensive series of calibrated resistances is required.

When registering photographically with a string galvanometer we have found it convenient to incorporate scale readings at the beginning and end of each record by moving the calibrated resistance, *C,* so as to produce a series of deflections of known resistance value in both directions from balance.

Behavior Research Fund
Chicago, Illinois

AN INTERVAL TIMER[1]

Franklin C. Paschal

The instrument here described was designed by the writer for the psychological laboratory at Vanderbilt University to fill the need of an automatic signalling device that could be readily adapted to a wide variety of situations. While there are instruments on the market that give excellent service in transmitting signals at regular intervals, every second or every 15 seconds, none are listed in the American catalogs that will give irregular signals without the construction of special parts for each particular use. The large number of combinations that may be obtained with this instrument and the ease with which it may be changed from one use to another will be evident from the explanation which follows.

A revolving brass arm, driven in a clockwise direction by a Telechron motor, makes one revolution per minute over the face of a 10-inch aluminum disk near the edge of which 120 holes have been bored. Pegs placed in any of these holes will make contact with the moving arm. The Telechron motor has a high degree of accuracy, since its rate of rotation is determined by the frequency of the alternating current and is not affected by variations in voltage. Almost all the large power companies now use master clocks to keep constant the rate of alternation of the current.

---

[1] The interval timer was constructed by Mr. M. Herblin of the instrument shops of the Vanderbilt University School of Medicine. Mr. Herblin is responsible for many of the refinements, including the contact lever and the pegs.

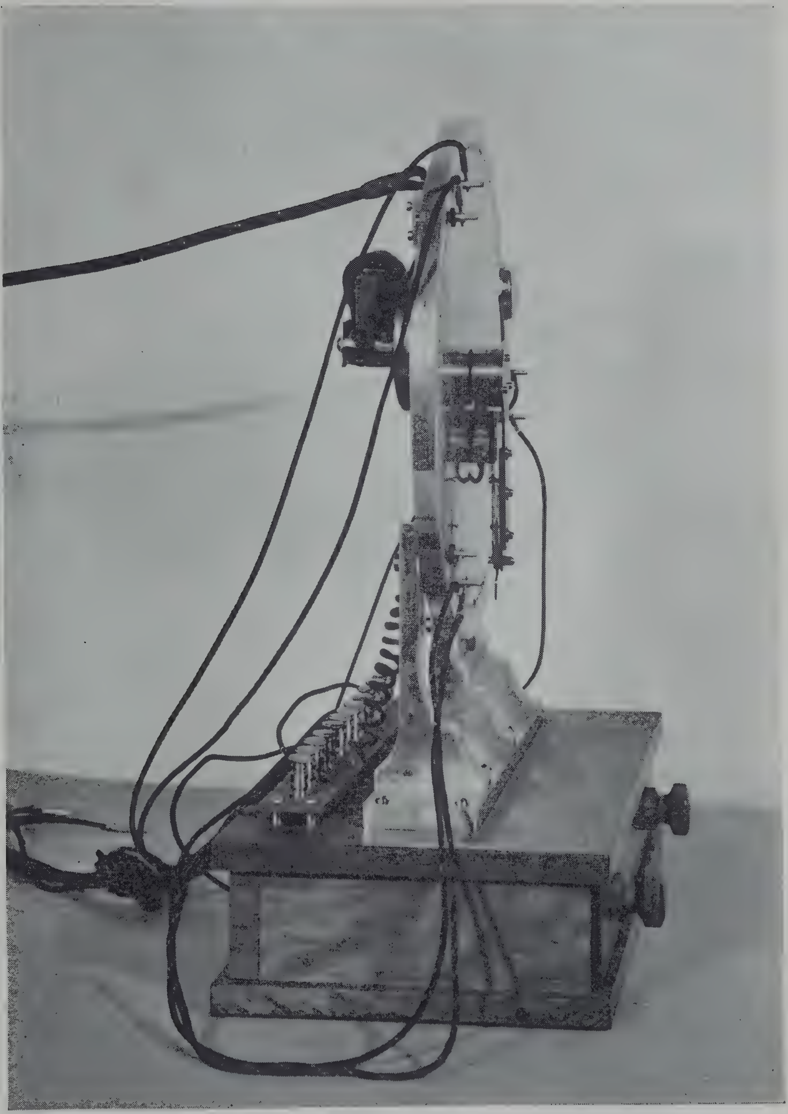


FIGURE 1

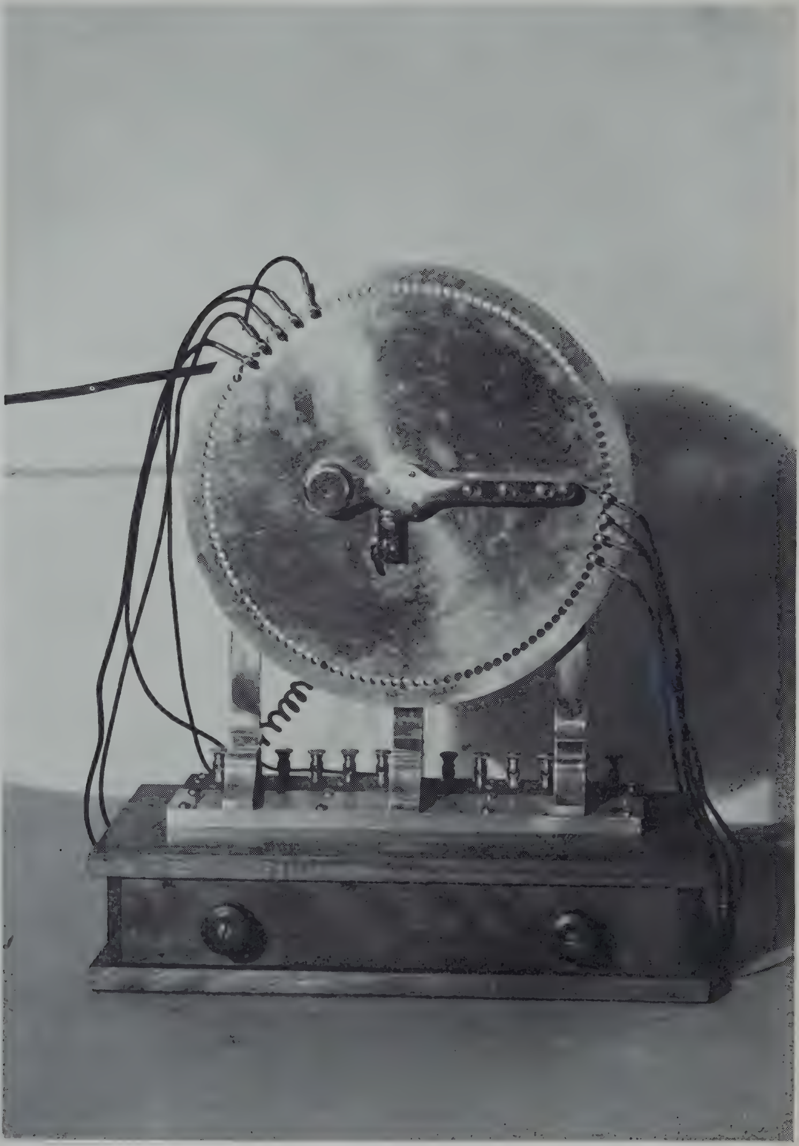


FIGURE 2

The moving arm has at its end a contact lever pivoted and held in position by a spring. The platinum point drags over the peg in passing and is instantly pulled into position by the spring without the back-lash or vibration that a flexible contact would have. Set-screws in the face of the armature permit adjustments of the contact point to compensate for wear or to lengthen or shorten the period of contact. The armature is carefully counterweighted and a spring exerts continuous pressure against the sleeve, preventing any free motion. This spring also serves as a means of electrical contact with the armature.

The small steel pegs have a hard rubber sheath which insulates them from the aluminum disk. A small jack is inserted into a lateral hole bored in the peg. On the rear of the base is found a row of binding posts, placed there to facilitate connecting the instrument externally. At the base of each binding-post is soldered a flexible wire 18 inches in length, which ends in the jack.

The side view of the interval timer (Figure 1) shows it as set up for a simple reaction-time experiment. The first contact of each pair gives the warning signals and the second contact which is set for two seconds later gives the stimulus signal. Three reactions per minute will be obtained from the set-up here shown. The front view (Figure 2) shows the instrument as set up for Dr. Woodrow's "detractor" experiment. The first contact sounds the warning signal, and the stimulus signal will be given at an interval of one or more seconds later depending upon the closing of one of the circuits outside the instrument. Two reactions per minute will be obtained.

A variety of motors and reduction gears are available in the Telechron line so that at comparatively small expense a different shaft speed can be provided making wider application possible. We would not choose this piece of apparatus, however, where intervals of less than a second were regularly required.

Vanderbilt University
Nashville, Tennessee

A PENDULUM CHRONOSCOPE

Starke Rosecrans Hathaway

In the use of a pendulum chronoscope there are two qualities which are of considerable importance. The first is that the recording and release mechanisms shall work silently when the stimulus is given. The second is that the readings obtained should be secured in some simple and permanent manner.

The present chronoscope attempts to meet these requirements. The operation of the release mechanism is silent, and there is but one moving part, the pendulum itself. Records are made on regular adding-machine paper tape. The essential part of the chronoscope is, of course, the pendulum. In considering various types of pendulums, we have attempted to select one which has a low air-resistance to its swing, and which would be as simple as possible to construct. The pendulum finally chosen is a single plate of brass shaped roughly like two folding fans connected at their centers (see *P*, Figure 1, and Figure 2). At this point of connection is the fulcrum. The upper part being somewhat smaller than the lower section gives the effect of a compound pendulum which is comparatively short in total length. Its period is slightly over one second.

Attached to one edge of the lower part is a small iron projection (*L*). This projection, when the pendulum is held in its initial position, fits against the face of a small electromagnet (*M*). The magnet has in its circuit a rheostat (*R*) which regulates the amount of current flowing. The magnet circuit includes the break key (*K*). When the break key is pressed, giving the stimulus, the magnet releases the pendulum which begins to fall from its initial position of 30 degrees from the vertical.

Attached to the middle of the lower part of the pendulum and projecting downward is a metal pointer (*N*). This swings just above a curved metal plate (*G*) so shaped as to follow accurately the locus of the pointer as it is carried by the pendulum. Between the pointer and the metal plate is a paper tape which runs from a roll at one end of the metal plate through guides following the curve of the plate to a cutter for tearing off the paper at the other end.

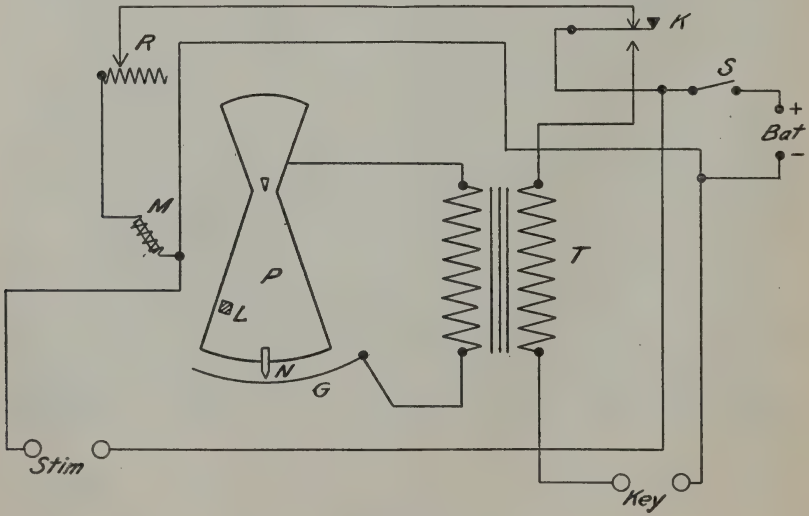


FIGURE 1

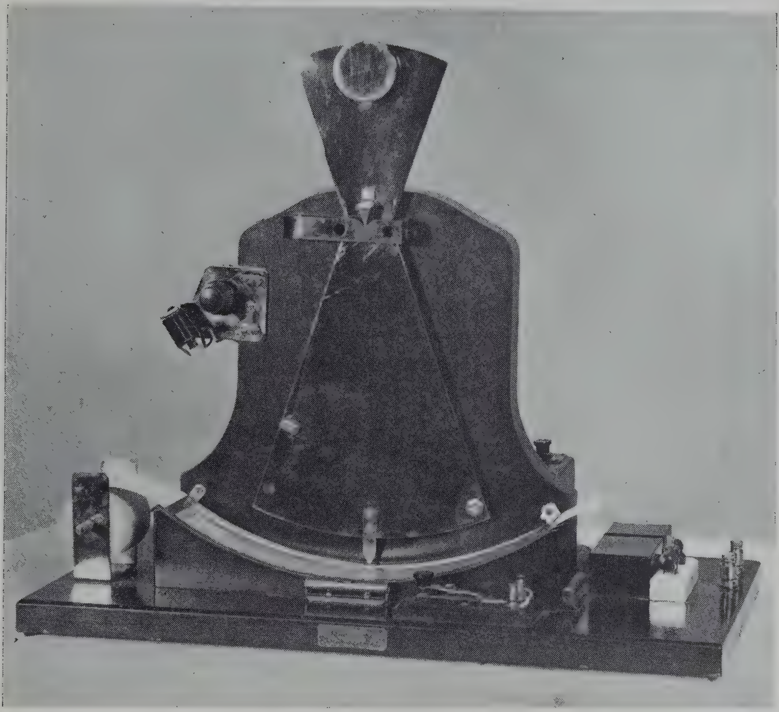


FIGURE 2

The secondary wires of a powerful induction coil (T) connect with the fulcrum of the pendulum and to the curved metal plate. This induction coil is activated by a 6-volt storage battery or four dry-cells, and, when the current flowing through its primary is broken, a spark passes from the pointer on the pendulum to the curved plate beneath, burning a hole through the paper tape. This hole marks the position of the pendulum at the moment of breaking the primary circuit.[1] The initial position of the pendulum when it is supported by the electromagnet is marked on the paper tape by use of a sharp pencil against a rule edge lying beneath the pointer at that position. The time of the reaction is shown by the distance between this initial position as marked by the pencil and the hole which is burned in the paper by the spark.

---

[1] Spark recording as a means of registering time has several well-known advantages. Professor Carl E. Seashore described his spark chronoscope in 1899 (1) and again, after eight years of continued use (2). The present simplified spark chronoscope is made by C. H. Stoelting Co., Chicago.

As assembled, the chronoscope includes the pendulum with its accessories, three pairs of binding posts, a key which is provided with a contact above for breaking the magnet and stimulus circuits, and one below for making the primary induction coil circuit. We may now describe a complete cycle in the functioning of the apparatus, and this will show more clearly how the parts are coordinated. The battery is connected to one set of binding posts. To the second set is connected the response key which is of the circuit-breaker type, and to the third set is attached the stimulus apparatus which is arranged so as to give the stimulus when the circuit is broken. A switch (*S*) closes the circuit through the magnet and through the binding posts which serve the stimulus apparatus. This switch is first closed and the pendulum is pushed up until it engages the supporting magnet. The rheostat is adjusted to the point where there is barely enough current to support the pendulum. This insures an almost immediate release of the pendulum from the magnet when the circuit is broken. The paper tape is now marked by pencil, indicating the position of the pointer, and the apparatus is ready for the first measurement. The before-mentioned break-make key, *K*, is now pressed, releasing the pendulum and presenting the stimulus. The make of the key prepares the circuit through the primary of the induction coil, to be closed by the subject when the response key attached to the second pair of binding posts is pressed. At the instant this key is pressed the spark from the secondary indicates the position of the pendulum by punching the hole in the paper tape. The paper tape is arranged so that the operator can pull it forward and tear off the strip on which the record has just been made. The apparatus is immediately made ready for the next measurement by simply replacing the pendulum, since the rheostat needs to be adjusted once only during the series of measurements.

In obtaining the time elapsing between the two points we have indicated on the strip of paper two methods are possible, (*a*) empirical calibration and (*b*) a mathematically calculated scale. First we may calibrate the pendulum in the usual manner, making a scale which will read directly in parts of a second the time represented by the distance the pendulum travels in a given time, as is indicated on the record. The second method is the one which has been adopted for reasons which will be indicated later. This method depends upon the fact that it is possible to calculate mathematically the time occupied by the pendulum in travelling from a definite initial position through any number of degrees or parts of a degree of its total swing. This calculation can be expressed in a table which will give the percentage of the total time period of the pendulum which has been consumed in traversing a given distance. The scale we would apply to our record in this case is a uniformly divided scale reading in parts of a degree of arc travelled by the pendulum from the initial position. We apply the scale to our record, read the number of degrees, and multiply the period of the pendulum by the percentage shown in our table representing that number of degrees on the scale.

Several important advantages are gained through use of this method: First, the cost of construction of the apparatus is decreased due to the fact that these percentages so represented hold for any pendulum which begins its swing from exactly 30 degrees from the vertical, thus doing away with the necessity of individual calibration of each pendulum, which is a quite laborious task. Secondly, this method allows for correct readings to be obtained, even though the pendulum period may vary through some accident or intentional lengthening or shortening or change in the pull of gravity. The readings are correct depending only upon the correct determination of the period of the pendulum with reasonable accuracy in the laboratory in which it is to be used. Thirdly, the method allows the experimenter to attach extra weights to the pendulum for the purpose of intentionally slowing up its period or of shortening it, as may be needed in some special problem, thus giving the apparatus increased flexibility for use. Without extra weights this chronoscope measures times up to .5 seconds. By adding a given pair of weights to the upper pendulum segment the available time is increased to 1.5 seconds.

REFERENCES

1. SEASHORE, C. E. New psychological apparatus. *Univ. Iowa Stud. Psychol.*, 1899, **2**, 153-158.

2. ————. The spark chronoscope. *Science*, 1907, **26**, 512-514.

Ohio University
Athens, Ohio

APPARATUS FOR A GROUP DEMONSTRATION OF WEBER'S LAW

G. C. GRINDLEY

The following is a description of a simple apparatus which may be used for demonstrating to a class Weber's law for intensity of illumination. The deviation from the law at low intensities may also be shown.

Figure 1 is a plan of the apparatus. A white screen, *SS*, such as that used with a lantern, is illuminated directly by an electric lamp, *L*, which has a bulb of white glass.

By means of a mirror, *M*, and a convex lens, *L'L'*, a bright image of an apperture, *A'A'*, is also thrown upon the screen. The aperture, *A'A'*, may conveniently be made in the form of a figure such as a cross, so that the class see a bright cross superimposed on the direct illumination of the screen. The illumination of the cross can be varied (without changing the direct illumination) by adjusting an iris diaphragm, *II*, in front of the lens. The adjustment of the iris diaphragm at which the cross is just perceptible can be found by any of the psychophysical methods.

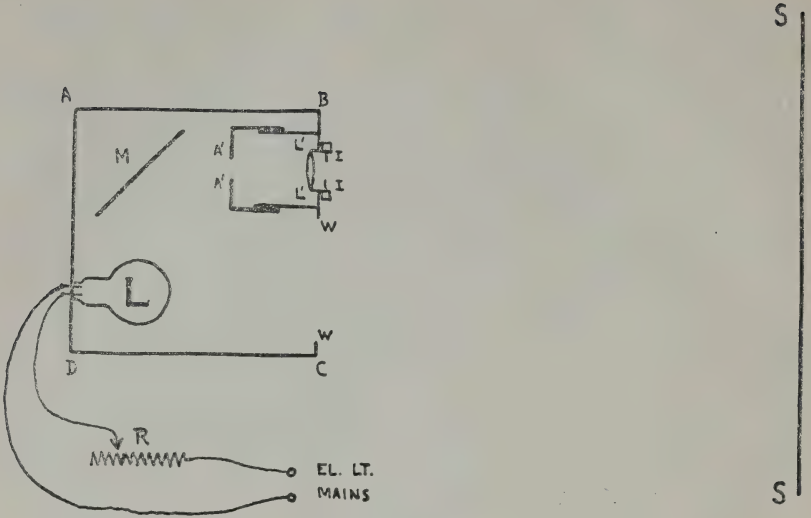


FIGURE 1
"Weber constant" $\delta I/I$
Illumination of screen (foot-candles)

The direct illumination of the screen is then altered by varying the resistance, *R*; and the adjustment of the iris diaphragm at which the cross is just perceptible is found for each value of the direct illumination. (It is clear that, if Weber's law is true, it will be found that the cross is just perceptible at the same adjustment of the iris diaphragm for all values of the direct illumination.)

In calibrating the apparatus the direct illumination of the screen for various values of the resistance, *R,* may easily be determined by the use of any photometric device such as an "illuminometer" or a "luxometer." The values of the differential threshold corresponding to various adjustments of the iris diaphragm may be calculated from similar measurements of the illumination of the cross when the brightness of the lamp is kept at a constant value and the window, *WW*, is covered with opaque material. These measurements should be made, and the results plotted as calibration curves, before demonstrating the experiment to a class.

In the actual demonstration, it is convenient to use the "limiting method." At each value of direct illumination the demonstrator increases (or decreases) the size of the iris diaphragm in a series of small steps, and at each step the members of the class indicate (e.g., by a show of hands) whether they can see the cross. The adjustment at which half the class think they see the cross is taken as the threshold value. It is convenient to plot the values of the "Weber constant," $\delta I/I$ (i.e., the values of the illumination of the cross divided by the direct illumination of the screen) against

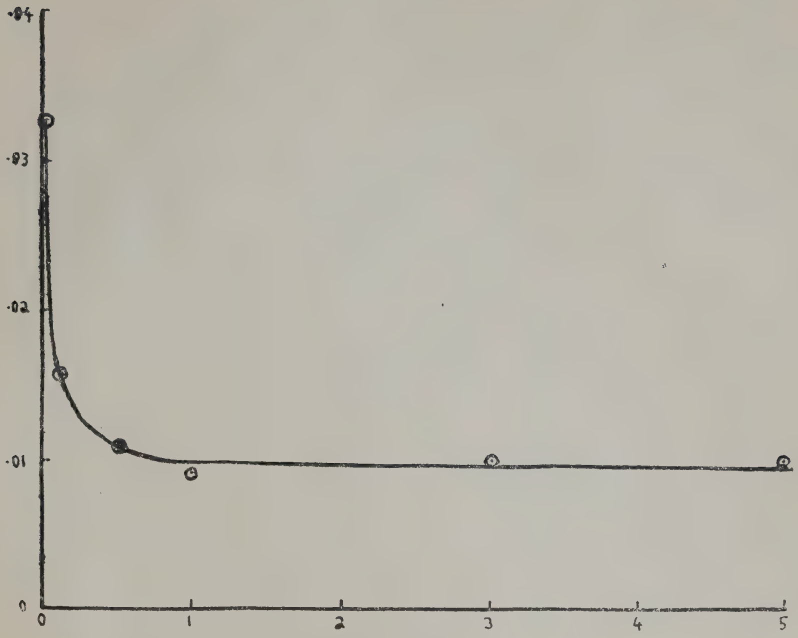


FIGURE 2

the direct illumination. Figure 2 is a typical curve obtained in this way with a class of 14 students. The whole experiment, including an explanation of the apparatus, took about half an hour.

Details of the Apparatus

In the apparatus at present in use in the Cambridge Laboratory the box, *ABCD*, is of three-ply wood, and measures 13″ x 13″ x 6″; the top may be removed to show the apparatus to the class. The lamp, *L*, is of 60 W. 105 V. The cross, *A′A′*, is cut in cardboard and may be moved backwards or forwards to focus the image on the screen. The arms of the cross are 1/2″ long and 1/10″ wide. The lens, *L′L′*, is of 3″ focal length and 1 1/2″ diameter. The iris diaphragm is that of a "Compur" photographic shutter, its maximum aperture being about equal to that of the lens. The position of the lever which varies the aperture of the diaphragm can be read by means of a scale attached to the front of the box.

The distance of the box from the screen is 4 feet, and at this distance the maximum direct illumination of the screen is about 5 foot-candles.

University of Cambridge
Cambridge, England

BOOKS

EDITED BY

KARL BÜHLER
W. J. CROZIER
JAMES DREVER
A. MICHOTTE
JOHN PAUL NAFE
A. L. SCHNIERMANN
RAYMOND H. WHEELER

KÖHLER, W. *Gestalt Psychology.* (3rd printing, revised.) New York: Liveright, 1929. Pp. x+402.

This book is the first extensive treatment of *Gestalttheorie* in which the contrast is sharply drawn between it and the "objective" theory of behaviorism, on the one hand, and the "subjective" theory of introspectionism, on the other. Some who have followed the experimental work of the *Gestalt* school with sympathetic interest have been inclined to accept those new results as an important addition to the existent corpus of experimental facts garnered by other methods. Others have been heard to say that they could find nothing "new" in the *Gestalt,* since they had always recognized a synthesis which is more than the sum of its analytic parts.

The thesis of Köhler's book corrects these misapprehensions and defends instead a dynamical theory of experience and behavior, which, though not unknown to the physicist, is radically opposed both to the "machine theory" of behaviorism and to the "meaning theory" of introspectionism.

Beginning with "direct experience," which is necessarily our point of departure in every investigation of facts or processes, "the simple truth," writes Köhler, "is that some of the experiences depending upon processes in my physical organism have the character of 'objectivity,' whereas others, depending upon other processes in the same organism, have a 'subjective' character, this contrast being altogether indifferent to the *genetic* 'subjectivity' of *both* types of processes and experiences, as depending upon the physical organism" (pp. 24 f.).

The problem of psychology cannot be solved, as the behaviorist thinks, by accepting the physical world naïvely, and rejecting the phenomena of direct experience. Psychology is too young a science to be able to solve its problems at once in quantitative terms. Instead, it must take note of the properties of direct experience, and construct from these the appropriate physiological processes which underlie experience. This should be no impossible task if, as we know to be the case, direct experience has allowed us to construct a physical world outside us, which is much less intimately related to experience than are the physiological events which underlie experience. If, for instance, I am able to experience one thing "between" two others, there must be something functional in the underlying nervous processes which corresponds to this "betweenness" (p. 64). Likewise "to a context experienced as 'one thing' belonging together, there corresponds a dynamical unit or whole in the underlying physiological processes" (p. 66).

No less "objective" than behaviorism is the direction of its interest upon "real events" rather than upon conscious states or appearances. *Gestalt* psychology does not neglect experience, neither does it interpret it as a mosaic of simple qualities. If it refutes behaviorism for its naïve "machine theory" of action, it is equally opposed to the "meaning theory" of the introspectionist, who counts his qualities as though they were indifferent bits of experience held together by an underlying physiology which must decide what shall be called "true" sensations (p. 88).

The argument is now directed against the machine theory which underlies both behaviorism and introspectionism. "Instead of reacting to local stimuli by local and mutually independent events, the organism reacts to an actual constellation of stimuli by a total process, which, as a functional whole, is its response to the whole situation" (p. 106).

In physical events, such as the play of electrons in an atom, or the movements of the stars in space, there is no "special topography laid down beforehand," such as both behaviorists and introspectionists assume to be the sole condition for response and experience. "Process is to be regarded as necessarily determined but among the various cases we find enormous differences in the relative influence which limiting topographical conditions, on the one hand, and the play of actual forces on the other, exert upon the course of events" (p. 112).

Writers on behaviorism and introspectionism alike seem never to "entertain the idea that some specific and orderly function might occur without being controlled either by special arrangements pre-established *ad hoc* or by arrangements acquired in learning" (p. 118). This third kind of orderly function is a *dynamical self-distribution* of forces in the sensorium which is controlled by conditions of the field as a whole, and not by a specific arrangement of nervous path-ways, either inherited or acquired.

Instead of the assumption attributed to James that there is a uniform continuity in original sensory experience, "and that all cuts and boundaries are introduced later on for pragmatic reasons" (p. 148), Köhler maintains that segregated units occur in the sensory field before they are named or otherwise qualified (p. 151). These units may or may not correspond to physical objects. It is wrong to suppose that objective *Gestalten* simply project themselves inside the organism. The segregation of wholes which we experience is determined within the nervous system. Waves of light, for instance, which mediate the sensory impression of an object seen "do not contain the slighest indication of any organization or any 'belonging together' which may exist among the parts of the objects by which they are reflected" (p. 175). Whenever we "attribute certain properties of sensory experience to the actual constellation of stimuli" (p. 176), we commit what may be called "the experience error." "As a matter of fact the whole retina is a mosaic of indifferently related spots and this is the case until sensory organization begins physiologically" (p. 177).

"Once we have realized, however, that stimulation, as such, is completely unorganized, the enormous biological value of sensory organization will become apparent. Since the rules governing this organization conform to the structure of objective units, to objective divisions, to objective 'belonging together,' in very many cases the result of their operation is a kind of reconstruction of those aspects of the objective physical situation which are temporarily lost on the way between the objects and the sense organs" (p. 171). In other words, topographical relationships constrain and promote a sensory organization which in some respects accords with the extended units or things of our physical environment. Thus our behavior, depending as it does upon the products of sensory organization, may at the same time be adequate to the objective nature of our environment.

Objective experience depends both on the things before us in the external world, and also on processes in the brain-field. But the spatial order of things as seen cannot be correlated with a merely geometrical localization of the corresponding processes in the brain-field. The structure of the brain will not lend itself to such a conception; instead we must assume a *dynamical contact* between neighboring areas of the brain. "Only with respect to those crudest properties of visual space, [such as] neighborhood and distance in general, will the underlying dynamical order of processes be extended homologously in brain-space" (p. 226). "At the same time 'my body' or 'myself' is a particular one among all the other sensory experiences, and, as such, it is also represented by definite particular processes in a definite part of the brain at a given moment. Dynamo-geometrical relations in the brain-field mean spatial relations in visual experiences. In this respect, other visual things are evidently outside or external to the special thing called 'my body.' But, you say, all these processes are *in* the brain, i.e., *in me,* after all! That does not matter for our problem, because experience depends upon the processes of a definite field, and experienced space depends upon the totality of dynamo-geometrical relations *in* it; the only meaning which 'in' and 'outside' have in experience is determined by this dynamical context, and to the physical place which this context has in the physical world, in the physical body, and more particularly in the physical skull, *there does not correspond any experience at all*" (p. 229).

The ascription to others of experiences similar to those we have ourselves is much more immediate than the inference that certain kinds of "bodies" have corresponding "minds" would allow. The similarity between one's direct sensory experience of another and this other's behavior is rooted in the dynamical interaction provoked by the other person's appearance in one's sensory field. The *crescendos* and *diminuendos* in the series of waves which his presence contributes to my sight and hearing are the more important bases for my understanding of him as an independent dynamic agent with motions like my own. "What surrounds us in sensory space, we describe as restless, or menacing, or friendly, because these are the dynamical attributes....which characterize the sensory experiences in question" (p. 264).

It would be impossible to develop a science of psychology from "consciousness" alone, because the field of actual experience always points beyond itself. It is these *directions* taken in experience which constitute its organization and account for the association of its outstanding features. Mere "neighborhood in space and time influences association only insofar as it determines organization" (p. 290). The theory of the conditioned reflex is inadequate because too little attention is paid to the conditions of organization, without which no association whatever can take place. "Our conclusion is," writes Köhler, "that association depends upon organization because association is just an after-effect of an organized process. Now, when we first introduced the concept of organization, we were hampered at every step by those empiristic hypotheses which sought to explain by the influence of meaning or other affiliated factors whatever experiences did not fit into the scheme of mosaic or machine theory. In many of these cases I hope that I have shown that fundamentally, the experience in question cannot be derived from previous learning in any form; and that, therefore, organization has to be accepted as an original aspect of our experiences. Now we may go further and make the statement that, on the contrary, wherever the influences of past experience can really be discovered in actual experience, this secondary effect itself is a consequence of past organization, because, generally, meaning will accompany an actual experience or be contained in it by a process of reproduction, and this reproduction will be based upon the trace of a larger organized whole" (pp. 299 f.).

With respect to social psychology, Köhler writes: "In *Gestalt* theory the varieties of directed attitudes are not considered as the operation of preexisting drives or instincts, but rather as the effects of actual situations" (pp. 324 f.). "It is difficult to understand, at the present time, how this directed attitude should depend upon the physiological situation of the organism as, for instance, hunger depends upon it. Nevertheless, for the most part, this attitude is quite similar to the need for food, and I do not hesitate to interpret it as a stress in the field between the self and those particular surrounding processes which are the physiological correlates of our experience of other persons" (p. 326).

As regards the much debated concept of "insight," to which Köhler devotes the final chapter of his book, the following citations will perhaps serve to state its meaning in *Gestalt* psychology:

Insight "does not mean more than our experience of definite determination in a context, an event or a development of the total field; and in the actual cases there need be nothing like an invention, or a new intelligent achievement....A total field would be experienced *without* insight, if all its several states, wholes, attitudes, etc., were simply given as a pattern, in which none was felt directly to depend upon any other and none to determine any other" (pp. 371 f.). "At the very outset it was proposed to use direct experience as an indicator of those physiological processes which, occurring between the realm of outer conditions, and the overt behavior of the organism, are not directly observable at the present time" (p. 373). "If....this particular dynamical context is taken to be the physiological fact underlying our experience of direct determination of attitude, no mystery remains in our description of common experience" (p. 375). "Instead of mystery, the most elementary concepts of physics are introduced into our view of the total physiological field. There is not the slightest reason why we should not experience determination *as such;* nor is there any reason why, in concrete cases, we should not feel just what our actual state or attitude is based upon; in one word, there is no reason why *insight* should not be an essential characteristic of much experience in the total field" (pp. 376 f.). "There is no mere sequence of indifferent events, connected indirectly. Each phase of what happens grows out of its predecessors, depending upon their concrete nature. And the subject whose experiences are an expression of this one developing context in the brain-field, will experience the development along with its 'referring to,' 'depending upon,' 'away from,' and so forth—that is, with *insight*."...."Even if learning should be the only way in which movements of the hand or of any other parts of the body become connected with the distributions of forces in the brain-field, the *right* movement will still be marked out by more than the final success it achieves" (pp. 390 f.).

R. M. Ogden

Cornell University
Ithaca, New York

Prince, M. *Clinical and Experimental Studies in Personality.* Cambridge, Mass.: Sci-Art Publishers, 1929. Pp. xvi+559.

The work of few men in any field of science is any more steadily accessible than is that of the late Morton Prince. A few well-known volumes present the substance of his major efforts. Recently Professor W. S. Taylor provided us with an outline-abstract of his contributions. And as a final gesture, he himself brought together this collection of twenty-one papers, which includes the best of his journal publications, so that we shall no longer be obliged to hunt the library files when we seek reference to his work.

There are included in this volume a dozen or more papers which appear never to have been published before, or, if so, in a very obscure manner. Some of these date as far back as twenty-five years ago, while others are quite recent. To these hitherto unpublished papers one turns, of course, with the greater interest, seeking to learn if perhaps they present anything importantly new; but one finds therein no more than the familiar arguments about the co-conscious and its place in the production of various psychopathic modes of behavior. Looking through them, as well as through the already familiar papers here reprinted, the reviewer was impressed perhaps as never before by the limited range of the literature with which Morton Prince appears to have worked. There are many references to his own publications, and relatively few to those of others. While he liked to speak the language of experimental psychology, especially of introspective psychology, there is a conspicuous absence of reference to the literature of introspective psychology, either experimental or theoretical. But the values and the weaknesses of the work of Morton Prince have often been pointed out. However they may be thought of, they are now made conveniently accessible and that is no mean contribution. The volume is thoroughly indexed, well printed, and serviceably bound, withal a good piece of workmanship. The heavy paper on which it is printed, however, makes it needlessly bulky.

EDMUND S. CONKLIN

University of Oregon
Eugene, Oregon

WHEELER, R. H. *The Science of Psychology.* New York: Crowell, 1929. Pp. xvii+556.

The *Science of Psychology* is the first attempt on the part of an American psychologist to give in textbook form a complete account of psychological facts in terms of Gestalt or Configuration.

One of the most revolutionary features of the book is its organization of material. Assuming that the whole is genetically and logically prior to its parts, "the book commences with the whole of which the human being psychologically is a part, namely, the social group." Thus the book reverses the conventional order of subject matter. A treatment of Social Behavior is followed by Intelligent Behavior, Emotive Behavior, Learning Behavior, Observational Behavior (including Simple Reactions, Space Perception, and the Finer Aspects of the Receptors), and a treatment of The Nervous System in its Relation to Behavior. Although a similar tendency in arrangement of subject matter is evident in other recent texts, the reversal is nowhere else as complete. In the light of the assumption that the whole is more than the sum of its parts, it is believed to be logical to begin the study of human behavior by a study of the social group, of which the individual is a part. The present writer, however, fails to see why this order is more logical than the conventional order. Biology, notably physiology, works from specialized parts to wholes, from the simple to the complex, from the undifferentiated to the differentiated. Granted that behavior-as-a-whole of the organism-as-a-whole should be studied with reference to society-as-a-whole. This, however, does not make it necessary that a student of psychology or of any other subject begin his science by working from the whole to its specialized parts. Knowledge is accumulative and proceeds from the relatively simple to the more complex. Why, then, should we do pedagogy violence by reversing its natural order for a study of psychology?

Another revolutionary feature of this book is its attempt to interpret (not explain, explanation is apparently unnecessary when we deal with *Gestalten*) all psychological facts in terms of the configurational point of view. The configurationalist points out that modern science does not regard the world as made up of composites or swarms of atom grains, but rather of an electrodynamic fabric, in which each component is an active determinant in the equilibrium form or motion of the total configuration. From this it follows that the whole cannot be described by its parts. Consequently the presuppositions upon which introspectionism and behaviorism rest are at fault. The introspectionist is at fault because he extracts only discrete elements from environment, and the behaviorist has been ensnared by the same misfeasance, for he accepts the same type of "machine theory" of the nervous system as does the introspectionist. The Gestalt school thus objects to analysis on the ground that it neglects certain over-all qualitative characteristics of complex experiences. But again it is difficult to see why an analysis in terms of parts should necessitate the neglect of any features of the whole. In general, the history of science suggests that analytical description is indispensable to intellectual progress. Dr. Wheeler, however, does not follow the Gestalt program to its logical conclusion. He actually does resort to analysis. He says, "There are two fundamental methods of science, the method of structural and functional analysis. The former yields descriptions and reduces the analyzed object or event to a lower or simpler order. The latter method yields explanation since it deals with those conditions which preserve the event. Data from the first end in classifications; those from the second lead to prediction and control." These methods, however, are incompatible with a strict adherence to the principles of Gestalt, insofar as the Gestalt school attempts to describe all human experience in terms of Gestalten. Furthermore, Dr. Wheeler resorts to a structural and functional analysis in spite of his own statement that "a whole disappears when we reduce it to its parts.....Now the human being as a unified, organized whole, exhibits such activities as seeing, feeling and thinking. It makes no difference whether or not we call these activities mental. As scientists we are not interested in the question whether these are activities of a mind. It is sufficient to know that they are activities of a human organism, as a whole, intact being, a being which cannot be explained in terms of its parts." The dilemma which has arisen is realized by Dr. Wheeler himself. Thus we read in the preface, "The text makes no distinction between the subjective and objective points of view as such, and stresses no dichotomy between existential and empirical psychology. It aims at a presentation free from philosophical presuppositions." If this really is the purpose of the book, why then does it attempt to explain all the facts of human experience in configurational terms? One may wonder if the varied phenomena of human nature can all be accounted for by a single theory which grew out of a study on the perception of movement. This strained condition becomes especially evident when the author tries to bring all the facts known about the learning process into obedience with the principles of configuration. According to the author, one should not, for example, be so old-fashioned as to recognize that anything of explanatory value may be derived from the facts concerning "trial and error" learning or "random movements." Instead, one should resort to such configurational ghosts as "insight," "maturation," "tensions," and "goals." Granted that learning does progress through insight, that maturation induces a change in forms of behavior, that goals are present in symbolic thought, and that tensions are set up whenever an animal is confronted with a stimulus. But if we wish to understand the laws which govern the learning process, or if, in terms of Gestalt, we wish to determine why one configuration emerges rather than another, it is not sufficient to refer merely to insight, goals, maturation, and tensions. To give the configurating faculty these names adds nothing. A similar condition exists in the various attempts to explain intelligent and emotive behavior. Whenever the configurationalist is in doubt or is pushed for analysis, it is comparatively easy for him to take refuge behind the word maturation, where, of course, he is safe.

Aside from such systematic difficulties as the author encounters, he offers a number of valuable contributions in this book. In the first place, the book is thoroughly up-to-date. The present writer knows of no other psychology text which offers as great a wealth of experimental material as this book. This material is presented in a very clear and readable manner. The central theme, namely, that based upon the principles of Gestalt and upon certain developments in physics, is followed very consistently throughout the book. A frequent restatement of these principles makes for clarity and aids the reader in keeping the author's purpose clearly in mind. Furthermore, this book gives a very clear analysis of introspection. Many new textbooks of psychology leave the student in total ignorance as to the method of introspection and its products. The book also gives a splendid analysis of classical methods and keen definitions of psychological terms. On the whole, the author shows a thorough acquaintance with historical and current psychological literature.

P. H. Ewert

University of Vermont
Burlington, Vermont

Marston, W. M. *Emotions of Normal People.* New York: Harcourt, Brace, 1928. Pp. xiii+405.

Analysis of the first quarter of this book serves to indicate the course the remainder is to take, for Marston has given us a systematic account of the emotional life. The first hundred pages lay the foundation upon which the superstructure is to be erected.

At the outset the psychology of emotion is defined, in a thoroughly conventional manner, as "the scientific description of affective consciousness." Conventionality ceases, however, when the first task of any system-maker, the definition of terms, is approached by the author. Consciousness is described as "the totality of energy generated within the junctional tissue between any two neurones, whenever the junctional membrane is continuously energized, from the emissive pole of one adjacent cell to the receptive pole of the next." The unit of conscious energy is the "psychon," analogous to the structural unit of neurology, the neuron. Psychonic energy generated in sensory synapses functions to give us sensation. Psychonic energy in motor synapses is also conscious! This is affective consciousness, which in its functioning is responsible for all affective and emotional experience. Whereas the term "sensation" is applicable to the consciousness that is sensory synaptic energy, "motation" is suggested as the proper name for motor consciousness.

Nothing new is to be found in the postulation of a motor consciousness and Marston claims no originality in this regard. He has turned to the defense and elaboration of a cause which science regarded as lost decades ago. The "innervation sense" has put behind it a long and shady past. The conception has hitherto been regarded as of doubtful value where, indeed, it has not aroused active antagonism. Enthusiasm for "feelings of innervation" is at a low ebb at present, to say the least.

Marston pictures the present situation in the psychology of emotion in this manner: "At the moment, investigators in the field of emotions find themselves at sea between the Scylla of James-Langeism and the Charybdis of youthful-minded adventurers in psychological research who would persuade us to hoist the Jolly Roger, abandon all theories and all previous results, and undertake statistical correlations of how all people react under all possible circumstances." To reach psychological *terra firma* we are directed to modify the physiography of Scylla. It is really a thing at which to marvel that that grand old rock, battered though it may be by persistent storms, still retains some semblance of its former self! It is stated that we may retain the original formulation of the James-Lange theory if we strike out all reference to perception of bodily changes. Marston maintains that emotion is "our awareness of our reaction *as it occurs*," and it is the last phrase that holds the significance of this statement. On the experiential side emotion is consciousness, the correlate of which in the nervous system is psychonic energy at the motor synapses.

The following are cited as "proofs" of the existence of motor consciousness: (1) Biologically, motor function is primary and sensory, and connector mechanisms are secondary. "It would be unexpected," the author states, "to find that the motor element....should itself completely fail of representation in the product, consciousness." (2) Motor neurons are distinguishable from sensory neurons with respect to structure and type of synaptic organization; the types of impulses passing over the two systems are different. What right have we to grant consciousness to one system and deny it to the other? (3) Certain motor phenomena may occur independently of sensory stimulation. How could such acts become conscious unless motor consciousness exists? (4) "Affective states accompanying motor discharge give every evidence of being far more diverse than the ensuing sensations of resulting bodily changes could possibly be." (5) "Affective tone may be changed by altering the motor set, without the slightest change in associated sensations." (6) The results of Head and Holmes are better explained on the basis of motor consciousness.

The first two and last are obviously "proofs" only to one who is already convinced of the existence of motor consciousness. Each of the remaining three involves questions of fact, and the facts are not clearly demonstrable. Certainly the possibilities in our modern conception of sensory integration are being overlooked in the fourth, and the fifth could be as easily and satisfactorily accounted for on the basis of an identification of the affective and the sensory. The author has failed to properly consider the latter possibility. It seems undeniable to many researchers in affection at present that feelings, on analysis, prove to be nothing more or less than patterned sensory experiences.

In discovering the "principles of primary feelings," Marston advances pleasantness and unpleasantness as the primary elements of motor consciousness. But he is careful to add that "primary element" does not mean "unit" from which all other feelings and emotions are compounded. Other energy forms must be added to produce the more complex emotions, just as a different and additional energy must be added to hydrogen and oxygen to form water. "Such complex, affective units, therefore, while all of them contain pleasantness or unpleasantness, or both (?),[1] must be expected to possess new attributes of motation which are actually not present in pleasantness and unpleasantness *per se*." Pleasantness, it is suggested, arises from the mutual facilitation of any two motor impulses on a motor psychon. Unpleasantness is then the result of antagonism between two or more motor impulses on a motor psychon. There is again nothing new or diverting in this conception. The view that pleasantness is due to harmonious combination and unpleasantness to discordant union is to be found, in the germ at least, in the writings of Johannes Müller, Herbart, and Aristotle. Marston's contribution in the use of this old idea lies in the correlation of these two states with various integrative relations brought about by combinations of tonic and phasic nervous impulses.

To establish the "primary emotions" the author finds it necessary to postulate a motor "self," which is identified with "continuous, tonic, motor discharge across motor psychons." "Motor stimuli" are *phasic* motor impulses which form psychonic conjunction with tonic motor excitations. The primary emotions are indicated by the various possibilities in the relations that may obtain between the motor self and motor stimuli, which, it is stressed, are not to be confused with environmental stimuli. "A primary emotion may be designated as an emotion which contains the maximal amount of alliance, antagonism, superiority of strength of the motor self in respect to the motor stimulus, or inferiority of strength of the motor self in respect to the motor stimulus." It is obvious that four primaries are possible. The following criteria were set in the selection of names that would designate the primaries: (1) The commonly understood meaning of the word must describe as completely as possible the relation between motor self and motor stimulus. (2) The name must suggest the experience in question. (3) Terms must be chosen which are not already "weighted with dissimilar affective meaning of literary origin." The suggested primaries, which were regarded as meeting these requirements, are: compliance, dominance, inducement, and submission. Other emotions are possible by the formation of different combinations of the motor self and motor stimuli. A scheme, not unlike the color circle in construction, has been devised by the author to represent all such possible relationships.

As has been indicated, the remainder of the book is largely the necessary consequence of the course charted up to this point. The primaries are considered at length, both with respect to descriptions of behavior involved in each and the mechanisms responsible for them. Combinations of the primaries are found to produce new integrative relations. Thus "passive compliance" and "active dominance" combine to form "desire"; "active compliance" unites with "passive dominance" to produce "satisfaction."

---

[1] The interrogation is mine.

Satisfaction and desire are capable of a new integration; a combination of these gives us "appetite."

One of the services of this book consists in exploding the notion that popular names of emotions are of prime significance to science. We have been trying for years to make literary names of emotions mean something to us and have retained them as bases of classification. Marston thinks that "what must be done is to give up attempts to define conflict-emotions, and go down to the very roots of biologically efficient behavior and discover the simple normal emotions that lie buried there."

The author has dealt less with normality of emotion, or rather abnormality, than one might expect from the title of the volume. The normative rather than the statistical definition of "normal" is adopted, which sets an obstacle in the way of the "hoisters of the Jolly Roger" who might otherwise be tempted to follow Marston in his systematic psychology of the emotions.

It is in the rôle of a systematic treatise that this book makes an appeal to the reviewer. He feels that it is in this light that it will be accepted or rejected by psychologists in general. If in giving the foregoing account he has caricatured the book in such a manner as to fail to lay due stress on the background of experiment and experience from which the latter chapters were in a large measure drawn, the reviewer asks the forgiveness of the author.

Frank A. Geldard

University of Virginia
University, Virginia

SUBJECT INDEX

Accommodation in binocular color vision, 301
Acoustics, records illustrating, 415, 416
Adaptation in binocular color vision, 298
Adaptive behavior of the eyes, 3
Affective factors in recall, 171
Ant, geotropism of, 102
Antagonism of reflexes and development of half-centers, 335
Anticipatory reactions in pursuit eye-movements, 3
Approximation and correction as adjustment pattern, 3
Arc cotangent function and hyperbola, 359
Auditory phenomena, records illustrating, 415, 416
Autocatalytic nature of growth and learning curve, 352

Behaviorism and modern psychology, 60
Bell Laboratory Victor records, 415, 416
Binocular color vision, quantitative study, 282
Blinking in binocular color vision, 302
Breakdown of eye-movement patterns, 3
Brightness as a condition of luminous appearance, 272
  of preexposure and color field, 15
  of surrounding field and color field, 15

Chronoscope, pendulum, 423
Color dominance, 282
  fields, shape and extent in relation to exposure conditions, 15
  mixture, 282
  vision, binocular, quantitative study, 282
  zones, 15
Color-weight illusion, 229
Competition of rival impulses for final common path and task resumption, 331
Conditioned reflex, time factors in, 372
  and hypnosis, 113

Death of Bekhterev, anniversary, 391
Difficulty of learning and length of problem, 157
Dominance, color, 282

Eidetic imagery in modern psychology, 58
Emotional reactions to strong stimuli, 133
Energy concept in modern psychology, 401
Equal appearing intervals, method of, empirical study, 192
Experimental method in psychology, 49
Eye-movements as adaptive behavior, 3
  in thinking of spatial relations, 379

Galvanic skin reflex and emotional reactions, 136
General emotionality theory (Burt), energy concept in, 403
Geotropism in ant, 102
Gestalt psychology as modern movement, 57

Half-centers and reflexes, 335
Hue in binocular color vision, 299
Human maze learning, manual guidance in, 213
Hyperbola and arc cotangent function, 357
Hypnosis and conditioned reflex, 113
  diminution of post-hypnotic suggestion, 366

I.E. ratio and emotional reactions, 149
Index of variability of *I.E.* ratio, 149
Inheritance of maze performance, 342
Inhibition and task resumption, 331
Insight, experimental study, 385
Instinct theory (McDougall), energy concept in, 402
Intensity in binocular color vision, 294
Interruption in work and tonus changes, 309

Interval timer, 420

Judging colored weights, vision and kinaesthesis in, 229

Kinaesthetic perception in judging colored weights, 229

Learning curves, mathematical study, 352, 359
  human, manual guidance in maze, 213
  length of problem and transfer, 154
  stylus-maze, effects of visual exposure, 36
Lectometer, description and reliability scores, 407
Length of problem and learning, 154
Luminosity and its varieties, 267
Luminous appearances, 247

Manual guidance in maze learning, 213
Mathematical study of learning curves, 352, 359
Maze behavior, inheritance, 342
  performance, factors, 338
Mean deviation of breath cycle, 147
    of pulse amplitude, 139
Memory, affective factors, 171
Mental work, completed and interrupted, tonus in, 309
Method of equal appearing intervals, empirical study, 192
  of paired comparisons, empirical study, 192
  right and wrong cases, 50
Mice, "spontaneous activity," 67
Mixture, color, 282
Modes of visual appearance, 247
Motivation and maze performance, 335

Neural tradition and psychic tension doctrine, 329

Ocular attitudes in thinking of spatial relations, 379
Order of merit method, empirical study, 192

Paired comparisons, method, empirical study, 192
Pendulum chronoscope, 423
Perceptual insight, experimental study, 385
Phi-gamma hypothesis, Thurstone's criticism, empirical study, 346
Philosophy and psychology, 49
Physics and psychology, 50
Physiology and psychology, 50
Post-hypnotic suggestion, measurement of diminution, 366
Practice effects, 163
"Psychic tensions," doctrine, 309
Psychophysical methods, empirical study, 191
Pulse in emotional reactions, 138

Rate of maze learning with visual exposure, 36
Rational learning, 154
Recall, effect of affective factors, 171
Reflex, antagonism in, and development of half-centers, 335
  conditioned, and hypnosis, 113
    time factors, 372
Reinforcement and task resumption, 330
Reliability of learning with visual exposure, 36
  of lectometer scores, 407
Respiration and emotional reaction, 141
Retention, length of problem, and transfer, 154
  and manual guidance, 221
Retinal rivalry, 282
Reward and maze performance, 338
Rivalry, 282

Saturation in binocular color vision, 296
Size of stimulus and color field, 15
"Spontaneous activity" of mice, 67
Subjective report of emotional reaction, 151

Temperature effects on activity of mice, 67
Thinking of spatial relations, ocular attitudes in, 379
Time estimation of "filled and unfilled" intervals, 391
    in hypnotic trance, 362
  factors in conditioned reflex, 372
Timer, interval, 420
Tonus in mental work, 309
Transfer and learning, 154
  practice effects, 163
Two-factor theory and energy concept, 402
    and modern psychology, 60

Uniocular fatigue in binocular color vision, 303

Vision, binocular, 282
Visual appearance, modes, classification, 248
  exposure in human maze learning, 36
  perception in judging colored weights, 229

Wave-length in binocular color vision, 287
Weber's law, apparatus for group demonstration, 427
Wheatstone-bridge method, continuous quantitative records by, 418
Will-temperament (Downey), energy concept in, 402

Zones, color, 15

468